

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/56, 15/80, 1/15, C12P 19/04, A01H 5/00, A61K 38/47, A23L 1/29		A1	(11) International Publication Number: WO 96/29416 (43) International Publication Date: 26 September 1996 (26.09.96)
(21) International Application Number: PCT/EP96/01009		(81) Designated States: AL, AM, AT, AU, AZ, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).	
(22) International Filing Date: 11 March 1996 (11.03.96)			
(30) Priority Data: 9505479.7 17 March 1995 (17.03.95) GB			
(71) Applicant (<i>for all designated States except US</i>): DANISCO A/S [DK/DK]; Langebrogade 1, P.O. Box 17, DK-1001 Copenhagen K (DK).		Published <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>	
(72) Inventors; and			
(75) Inventors/Applicants (<i>for US only</i>): MADRID, Susan, Mampusta [DK/DK]; Ellefolden 7, Trørød, DK-2950 Vedbaek (DK). RASMUSSEN, Preben [DK/DK]; Muldvad 45, DK-2800 Lyngby (DK). BARUCH, Anita [DK/DK]; Brøndbyvestervej 9, DK-2600 Glostrup (DK).			
(74) Agents: HARDING, Charles, Thomas et al.; D. Young & Co., 21 New Fetter Lane, London EC4A 1DA (GB).			

(54) Title: **ASPERGILLUS ARABINOFURANOSIDASE**

(57) Abstract

An enzyme capable of degrading arabinoxylan is described. In addition, there is described a nucleotide sequence coding for the same and a promoter for controlling the expression of the same.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgyzstan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic	KR	Republic of Korea	SE	Sweden
CG	Congo	KZ	Kazakhstan	SG	Singapore
CH	Switzerland	L1	Liechtenstein	SI	Slovenia
CI	Côte d'Ivoire	LK	Sri Lanka	SK	Slovakia
CM	Cameroon	LR	Liberia	SN	Senegal
CN	China	LT	Lithuania	SZ	Swaziland
CS	Czechoslovakia	LU	Luxembourg	TD	Chad
CZ	Czech Republic	LV	Latvia	TG	Togo
DE	Germany	MC	Monaco	TJ	Tajikistan
DK	Denmark	MD	Republic of Moldova	TT	Trinidad and Tobago
EE	Estonia	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	UG	Uganda
FI	Finland	MN	Mongolia	US	United States of America
FR	France	MR	Mauritania	UZ	Uzbekistan
GA	Gabon			VN	Viet Nam

ASPERGILLUS ARABINOFURANOSIDASE

The present invention relates to an enzyme. In addition, the present invention relates to a nucleotide sequence coding for the enzyme. Also, the present invention relates to a 5 promoter, wherein the promoter can be used to control the expression of the nucleotide sequence coding for the enzyme.

In particular, the enzyme of the present invention is an arabinofuranosidase enzyme having arabinoxylan degrading activity.

10

It is known that it is desirable to direct expression of a gene of interest ("GOI") in certain tissues of an organism - such as a filamentous fungus (such as *Aspergillus Niger*) or even a plant crop. The resultant protein or enzyme may be useful for the organism itself. For example, it may be desirable to produce crop protein products with an optimised amino 15 acid composition and so increase the nutritive value of a crop. For example, the crop may be made more useful as a feed.

20

In the alternative, it may be desirable to isolate the resultant protein or enzyme and then use the protein or enzyme to prepare, for example, food compositions. In this regard, the resultant protein or enzyme can be a component of the food composition or it can be used to prepare food compositions, including altering the characteristics or appearance of food compositions. It may even be desirable to use the organism, such as a filamentous fungus or a crop plant, to express non-plant genes, such as for the same purposes.

25

Also, it may be desirable to use an organism, such as a filamentous fungus or a crop plant, to express mammalian genes. Examples of the latter products include interferons, insulin, blood factors and plasminogen activators. It is also desirable to use micro-organisms, such as filamentous fungi, to prepare products from GOIs by use of promoters 30 that are active in the micro-organisms.

Fruit and vegetable cell walls largely consist of polysaccharide, the major components being pectin, cellulose and xyloglucan (R.R. Selvendran and J.A. Robertson, IFR Report 1989). Numerous cell wall models have been proposed which attempt to incorporate the essential properties of strength and flexibility (P. Albersheim, Sci. Am. 232, 81-95.

- 5 1975; P. Albersheim, Plant Biochem. 3rd Edition (Bonner and Varner), Ac. Press, 1976; T. Hayashi, Ann. Rev. Plant Physiol. & Plant Mol. Biol., 40, 139-168, 1989).

The composition of the plant cell wall is complex and variable. Polysaccharides are mainly found in the form of long chains of cellulose (the main structural component of 10 the plant cell wall), hemicellulose (comprising various β -xylan chains) and pectic substances (consisting of galacturonans and rhamnogalacturonans; arabinans; and galactans and arabinogalactans). From the standpoint of the food industry, the pectic substances, arabinans in particular, have become one of the most important constituents of plant cell walls (Whitaker, J.R. (1984) Enzyme Microb. Technol., 6,341).

15

One form of plant polysaccharide is arabinan. A review of arabinans may be found in EP-A-0506190. According to this document, arabinans consist of a main chain of α -(1 \rightarrow 5) groups linked to one another. Side chains are linked α -(1 \rightarrow 3) or sometimes α -(1 \rightarrow 2) to the main α -(1 \rightarrow 5)-L-arabinan backbone. In apple, for example, one third of the 20 total arabinose is present in the side chains. The molecular weight of arabinan is normally about 15 kDa.

Arabinans are degraded by enzymes collectively called arabinases. In this regard, arabinan-degrading activity is the ability of an enzyme to release arabinose residues, either monomers or oligomers, from arabinan backbones or from arabinan-containing side chains of other hemicellulose backbone structures such as arabinogalactans, or even the release of arabinose monomers via the cleavage of the 1 \rightarrow 6 linkage between the terminal arabinofuranosyl unit and the intermediate glucosyl unit of monoterpenyl α -L-arabinofuranosyl glucosides.

30

The activity of the arabinan degrading enzymes of EP-A-0506190 include: a) the ability to cleave (1 \rightarrow 2)- α -L-arabinosidic linkages; b) the ability to cleave (1 \rightarrow 3)- α -L-arabinosidic linkages; c) the ability to cleave (1 \rightarrow 5)- α -L-arabinosidic linkages; d) the ability to cleave the 1 \rightarrow 6 linkage between the terminal arabinofuranosyl unit and the intermediate glucosyl 5 unit of monoterpenyl α -L-arabinofuranosyl glucosides.

Arabinan-degrading enzymes are known to be produced by a variety of plants and microorganisms, among these, fungi such as those of the genera *Aspergillus*, *Corticium*, *Rhodotorula* (Kaji, A. (1984) Adv. Carbohydr. Chem. Biochem., 42, 383), *Dichotomitus* 10 (Brillouet et al. (1985) Carbohydrate Research, 144, 113), *Ascomycetes* and *Basidiomycetes* (Sydow, G. (1977) DDR Patent Application No. 124.812).

Another plant polysaccharide is xylan, whose major monosaccharide unit is xylose. Xylans are abundant components of the hemicelluloses. In monocotyledonous plants the 15 dominant hemicellulose is an arabinoxylan, in which arabinose side chains are attached to a backbone of xylose residues.

Arabinoxylans are carbohydrates found in the cell wall of cereals. A review of arabinoxylans and the enzymatic degradation thereof may be found in Voragen *et al* 20 (1992 Characterisation of Cereal Arabinoxylans, Xylans and Xylanases pages 51-67, edited by J. Visser published by Elsevier Science Publishers).

Typically, arabinoxylans comprise a xylose backbone linked together via β -1,4- bonds. The xylose backbone is substituted with L-arabinose residues which are linked via α -1 25 bonds to the 2 or 3 position of the xylose residues. The xylose residues can be single or double substituted. In addition to substitution with arabinose the xylose residues can be substituted with acetyl groups, glucuronic acid and various other carbohydrates. The arabinose residues can be further substituted with phenolic acids such as ferulic acid and coumaric acid. The degree and kind of substitution depends on the source of the 30 particular arabinoxylan.

Arabinoxylans are found in cereal cell wall where they are part of the secondary cell wall. Arabinoxylans form about 3 % of wheat flour - part of it is water soluble (WSP), part of it is water insoluble (WIP).

- 5 Despite the fact that the arabinoxylans amount to only about 3 % of wheat the importance of the arabinoxylan fraction is much higher. This is because the arabinoxylans of cereals act as hydrocolloids, as they form a gel like structure with water. For example, the arabinoxylans of wheat flour bind up to 30% of the water in a dough despite the fact that they amount to only 3 % of the dry matter. When arabinoxylans bind water they
10 increase the viscosity of the ground cereals and to such an extent that the cereals can become difficult to manage.

The rheological properties of several systems where ground cereals are used can be manipulated using enzymes that degrade arabinoxylans. In modern bakery it is
15 advantageous to reduce the viscosity of the dough in order to reduce the energy needed to process the doughs and also to get a higher volume of the bread. This is usually achieved by using enzymes that can degrade the xylose backbone of arabinoxylans.

Enzymes that only cleave the arabinose side chains from the xylan backbone of
20 arabinoxylan are, for the purposes of this application, collectively called arabinoxylan degrading enzymes.

In feeds based on cereals, arabinoxylans in the cereals can increase the viscosity of the fluids in the intestines of the animals after the feeds have been ingested. This is a
25 problem as it causes discomfort, such as indigestion, to the animals. Also, the nutritive value of the feeds is reduced. These problems can be avoided by addition of enzymes that degrade the arabinoxylan (such as xylanases) to the feed to avoid indigestion and to increase the nutritive value of the feed. However, some enzymes that degrade the arabinoxylans (especially some of the xylanases) require the presence of unsubstituted
30 backbones and so their activity can be limited.

Further discussions on arabinoxylans can be found in *Xylans and Xylanases* (1992, edited by J. Visser published by Elsevier Science Publishers).

An arabinoxylan degrading enzyme is (1,4)- β -D-arabinoxylan arabinofuranohydrolase (AXH), as described by Kormelink *et al* 1991 (Kormelink, F.J.M., Searle-Van Leeuwen M.J.F., Wood, T.M., Voragen, A.G.J.(1991) Purification and characterization of a (1,4)- β -D-arabinoxylan arabinofuranohydrolase from *Aspergillus awamori*. Appl. Microbiol. Biotechnol. 25:753-758). However, this document provides no sequence data for the enzyme or the nucleotide sequence coding for same or for the promoter for the same.

10

Clearly, it would be useful to be able to degrade arabinoxylans, preferably by use of recombinant DNA techniques.

The present invention seeks to provide an enzyme having arabinoxylan degrading activity; 15 preferably wherein the enzyme can be prepared in certain or specific cells or tissues, such as in just a specific cell or tissue, of an organism, typically a filamentous fungus, preferably of the genus *Aspergillus*, such as *Aspergillus niger*, or even a plant.

Also, the present invention seeks to provide a GOI coding for the enzyme that can be 20 expressed preferably in specific cells or tissues, such as in certain or specific cells or tissues, of an organism, typically a filamentous fungus, preferably of the genus *Aspergillus*, such as *Aspergillus niger*, or even a plant.

In addition, the present invention seeks to provide a promoter that is capable of directing 25 expression of a GOI, such as a nucleotide sequence coding for the enzyme according to the present invention, preferably in certain specific cells or tissues, such as in just a specific cell or tissue, of an organism, typically a filamentous fungus, preferably of the genus *Aspergillus*, such as *Aspergillus niger*, or even a plant. Preferably, the promoter is used in *Aspergillus* wherein the product encoded by the GOI is excreted from the host 30 organism into the surrounding medium.

Furthermore, the present invention seeks to provide constructs, vectors, plasmids, cells, tissues, organs and organisms comprising the GOI and/or the promoter, and methods of expressing the same, preferably in specific cells or tissues, such as expression in just a specific cell or tissue, of an organism, typically a filamentous fungus, preferably of the 5 genus *Aspergillus*, or even a plant.

According to a first aspect of the present invention there is provided an enzyme obtainable from *Aspergillus*, wherein the enzyme has the following characteristics: a MW of 33,270 D ± 50 D; a pI value of about 3.7; arabinoxylan degrading activity; a pH 10 optima of from about 2.5 to about 7.0 (more especially from about 3.3 to about 4.6, more especially about 4); a temperature optima of from about 40°C to about 60°C (more especially from about 45°C to about 55°C, more especially about 50°C); and wherein the enzyme is capable of cleaving arabinose from the xylose backbone of an arabinoxylan.

15 According to a second aspect of the present invention there is provided an enzyme having the sequence shown as SEQ. I.D. No. 1 or a variant, homologue or fragment thereof.

According to a third aspect of the present invention there is provided an enzyme coded by the nucleotide sequence shown as SEQ. I.D. No. 2 or a variant, homologue or 20 fragment thereof or a sequence complementary thereto.

According to a fourth aspect of the present invention there is provided a nucleotide sequence coding for the enzyme according to the present invention.

25 According to a fifth aspect of the present invention there is provided a nucleotide sequence having the sequence shown as SEQ. I.D. No. 2 or a variant, homologue or fragment thereof or a sequence complementary thereto.

According to a sixth aspect of the present invention there is provided a promoter having 30 the sequence shown as SEQ. I.D. No. 3 or a variant, homologue or fragment thereof or a sequence complementary thereto.

According to a seventh aspect of the present invention there is provided a terminator having the nucleotide sequence shown as SEQ. I.D. No. 13 or a variant, homologue or fragment thereof or a sequence complementary thereto.

5 According to an eighth aspect of the present invention there is provided a signal sequence having the nucleotide sequence shown as SEQ. I.D. No. 14 or a variant, homologue or fragment thereof or a sequence complementary thereto.

10 According to a ninth aspect of the present invention there is provided a process for expressing a GOI by use of a promoter, wherein the promoter is the promoter according to the present invention.

15 According to a tenth aspect of the present invention there is provided the use of an enzyme according to the present invention to degrade an arabinoxylan.

According to an eleventh aspect of the present invention there is provided a combination of enzymes to degrade an arabinoxylan, the combination comprising an enzyme according to the present invention and a xylanase.

20 According to a twelfth aspect of the present invention there is provided plasmid NCIMB 40703, or a nucleotide sequence obtainable therefrom for expressing an enzyme capable of degrading arabinoxylan or for controlling the expression thereof or for controlling the expression of another GOI.

25 According to a thirteenth aspect of the present invention there is provided a signal sequence having the sequence shown as SEQ. I.D. No. 15 or a variant, homologue or fragment thereof.

30 According to a fourteenth aspect of the present invention there is provided the use of the enzyme according to the present invention in the manufacture of a medicament or foodstuff to reduce or prevent indigestion and/or increase digestibility and/or increase nutrient absorption.

According to a fifteenth aspect of the present invention there is provided an arabinofuranosidase enzyme having arabinoxylan degrading activity, which is immunologically reactive with an antibody raised against a purified arabinofuranosidase enzyme having the sequence shown as SEQ. I.D. No. 1.

5

According to a sixteenth aspect of the present invention there is provided an arabinofuranosidase promoter wherein the promoter is inducible by an intermediate in xylose metabolism.

10 According to a seventeenth aspect of the present invention there is provided a process of reducing the viscosity of a branched substrate wherein the enzyme degrades the branches of the substrate but not the backbone of the substrate.

15 According to a further aspect of the present invention there is provided the use of the enzyme of the present invention as a viscosity modifier.

According to a further aspect of the present invention there is provided the use of the enzyme of the present invention to reduce the viscosity of pectin.

20 Other aspects of the present invention include constructs, vectors, plasmids, cells, tissues, organs and transgenic organisms comprising the aforementioned aspects of the present invention.

25 Other aspects of the present invention include methods of expressing or allowing expression or transforming any one of the nucleotide sequence, the construct, the plasmid, the vector, the cell, the tissue, the organ or the organism, as well as the products thereof.

30 Additional aspects of the present invention include uses of the promoter for expressing GOIs in culture media such as a broth or in a transgenic organism.

Further aspects of the present invention include uses of the enzyme for preparing or treating foodstuffs, including animal feed.

Preferably the enzyme is coded by the nucleotide sequence shown as SEQ. I.D. No. 2
5 or a variant, homologue or fragment thereof or a sequence complementary thereto.

Preferably the nucleotide sequence has the sequence shown as SEQ. I.D. No. 2 or a
variant, homologue or fragment thereof or a sequence complementary thereto.

10 Preferably the nucleotide sequence is operatively linked to a promoter.

Preferably the promoter comprises the sequence CCAAT.

Preferably the promoter is the promoter having the sequence shown as SEQ. I.D. No.
15 3 or a variant, homologue or fragment thereof or a sequence complementary thereto.

Preferably, the promoter comprises the 100 bps sequence from the *Xma* 111 to the
*BamH*1 sites.

20 Preferably the promoter of the present invention is operatively linked to a GOI.

Preferably the GOI comprises a nucleotide sequence according to the present invention.

Preferably the transgenic organism is a fungus.

25

Preferably the transgenic organism is a filamentous fungus, more preferably of the genus
Aspergillus.

Preferably the transgenic organism is a plant.

30

Preferably, in the use, the enzyme is used in combination with a xylanase, preferably an
endoxylanase.

Highly preferred embodiments of each of the aspects of the present invention do not include any one of the native enzyme, the native promoter or the native nucleotide sequence in its natural environment.

5 Preferably, in any one of the plasmid, the vector such as an expression vector or a transformation vector, the cell, the tissue, the organ, the organism or the transgenic organism, the promoter is present in combination with at least one GOI.

10 Preferably the promoter and the GOI are stably incorporated within the transgenic 10 organism's genome.

Preferably the transgenic organism is a filamentous fungus, preferably of the genus *Aspergillus*, more preferably *Aspergillus niger*. The transgenic organism can even be a plant, such as a monocot or dicot plant.

15 A highly preferred embodiment is an enzyme obtainable from *Aspergillus*, wherein the enzyme has the following characteristics: a MW of 33,270 D \pm 50 D; a pI value of about 3.7; arabinoxylan degrading activity; a pH optima of from about 2.5 to about 7.0 (more especially from about 3.3 to about 4.6, more especially about 4); a temperature 20 optima of from about 40°C to about 60°C (more especially from about 45°C to about 55°C, more especially about 50°C); and wherein the enzyme is capable of cleaving arabinose from the xylose backbone of an arabinoxylan; wherein the enzyme has the sequence shown as SEQ. I.D. No. 1 or a variant, homologue or fragment thereof.

25 Another highly preferred embodiment is an enzyme obtainable from *Aspergillus*, wherein the enzyme has the following characteristics: a MW of 33,270 D \pm 50 D; a pI value of about 3.7; arabinoxylan degrading activity; a pH optima of from about 2.5 to about 7.0 (more especially from about 3.3 to about 4.6, more especially about 4); a temperature optima of from about 40°C to about 60°C (more especially from about 45°C to about 30 55°C, more especially about 50°C); and wherein the enzyme is capable of cleaving arabinose from the xylose backbone of an arabinoxylan; wherein the enzyme is coded by the nucleotide sequence shown as SEQ. I.D. No. 2 or a variant, homologue or fragment

thereof or a sequence complementary thereto.

The advantages of the present invention are that it provides a means for preparing an arabinofuranosidase enzyme having arabinoxylan degrading activity and the nucleotide 5 sequence coding for the same. In addition, it provides a promoter that can control the expression of that, or another, nucleotide sequence.

Other advantages are that the enzyme of the present invention can affect the viscosity of ground cereals, such as dough, to ease the handling thereof and for example to get a 10 higher volume of the bread.

The enzyme of the present invention is also advantageous for feed because it degrades arabinoxylan and thus increases the nutritive value of the feed. In addition, it reduces the viscosity of the arabinoxylan in the intestine of the animals and so reduces or prevents 15 indigestion.

The combination of the use of the enzyme of the present invention with a xylanase is particularly advantageous because the enzyme of the present invention and the xylanase have a surprising and unexpected synergistic effect with each other.

20

In this regard, the enzyme of the present invention increases the degradative effect of the xylanase, and the xylanase increases the degradative effect of the enzyme of the present invention. It is believed that the activity of the xylanase is increased because the enzyme 25 of the present invention provides a polysaccharide substrate having fewer substituted groups.

The present invention therefore provides an enzyme having arabinoxylan degrading activity wherein the enzyme can be prepared in certain or specific cells or tissues, such as in just a specific cell or tissue, of an organism, typically a filamentous fungus, 30 preferably of the genus *Aspergillus*, such as *Aspergillus niger*. The enzyme may even be prepared in a plant.

More in particular, the enzyme of the present invention is capable of specifically cleaving arabinose from the xylose backbone of arabinoxylan.

The 5 arabinofuranosidase of the present invention is different from the arabinofuranosidases previously known. In this regard, the previous described arabinofuranosidases - such as those of EP-A-0506190 - are characterised by their ability to degrade unbranched arabinan, and are assayed using p-nitrophenyl-arabinoside.

The 10 arabinofuranosidase of the present invention does not degrade unbranched arabinan, and only a minor activity is seen on nitrophenyl-arabinoside. In contrast, the arabinofuranosidase of the present invention is useful for degrading arabinoxylan. Therefore, the arabinofuranosidase of the present invention is quite different from the previous isolated arabinofuranosidases.

15 Also, the present invention provides a GOI coding for the enzyme that can be expressed preferably in specific cells or tissues, such as in certain or specific cells or tissues, of an organism, typically a filamentous fungus, preferably of the genus *Aspergillus*, such as *Aspergillus niger*. The GOI may even be expressed in a plant.

20 In addition, the present invention provides a promoter that is capable of directing expression of a GOI, such as a nucleotide sequence coding for the enzyme according to the present invention, preferably in certain specific cells or tissues, such as in just a specific cell or tissue, of an organism, typically a filamentous fungus, preferably of the genus *Aspergillus*, such as *Aspergillus niger*, or even a plant. Preferably, the promoter 25 is used in *Aspergillus* wherein the product encoded by the GOI is excreted from the host organism into the surrounding medium. The promoter may even be tailored (if necessary) to express a GOI in a plant.

30 The present invention also provides constructs, vectors, plasmids, cells, tissues, organs and organisms comprising the GOI and/or the promoter, and methods of expressing the same, preferably in specific cells or tissues, such as expression in just a specific cell or tissue, of an organism, typically a filamentous fungus, preferably of the genus

Aspergillus, or even a plant.

- The terms "variant", "homologue" or "fragment" in relation to the enzyme include any substitution of, variation of, modification of, replacement of, deletion of or addition of 5 one (or more) amino acid from or to the sequence providing the resultant amino acid sequence has arabinoxylan degrading activity, preferably having at least the same activity of the enzyme shown in the sequence listings (SEQ I.D. No. 1 or 12). In particular, the term "homologue" covers homology with respect to structure and/or function providing the resultant enzyme has arabinoxylan degrading activity. With respect to sequence 10 homology, preferably there is at least 75%, more preferably at least 85%, more preferably at least 90% homology to SEQ ID NO. 1 shown in the attached sequence listings. More preferably there is at least 95%, more preferably at least 98%, homology to SEQ ID NO. 1 shown in the attached sequence listings.
- 15 The terms "variant", "homologue" or "fragment" in relation to the nucleotide sequence coding for the enzyme include any substitution of, variation of, modification of, replacement of, deletion of or addition of one (or more) nucleic acid from or to the sequence providing the resultant nucleotide sequence codes for an enzyme having arabinoxylan degrading activity, preferably having at least the same activity of the 20 enzyme shown in the sequence listings (SEQ I.D. No. 2 or 12). In particular, the term "homologue" covers homology with respect to structure and/or function providing the resultant nucleotide sequence codes for an enzyme having arabinoxylan degrading activity. With respect to sequence homology, preferably there is at least 75%, more preferably at least 85%, more preferably at least 90% homology to SEQ ID NO. 2 shown 25 in the attached sequence listings. More preferably there is at least 95%, more preferably at least 98%, homology to SEQ ID NO. 2 shown in the attached sequence listings.

The terms "variant", "homologue" or "fragment" in relation to the promoter include any substitution of, variation of, modification of, replacement of, deletion of or addition of 30 one (or more) nucleic acid from or to the sequence providing the resultant nucleotide sequence has the ability to act as a promoter in an expression system - such as the transformed cell or the transgenic organism according to the present invention. In

particular, the term "homologue" covers homology with respect to structure and/or function providing the resultant nucleotide sequence has the ability to act as a promoter. With respect to sequence homology, preferably there is at least 75%, more preferably at least 85%, more preferably at least 90% homology to SEQ ID NO. 3 shown in the attached sequence listings. More preferably there is at least 95%, more preferably at least 98%, homology to SEQ ID NO. 3 shown in the attached sequence listings.

The terms "variant", "homologue" or "fragment" in relation to the terminator or signal nucleotide sequences include any substitution of, variation of, modification of, replacement of, deletion of or addition of one (or more) nucleic acid from or to the sequence providing the resultant nucleotide sequence has the ability to act as a terminator or codes for an amino acid sequence that has the ability to act as a signal sequence respectively in an expression system - such as the transformed cell or the transgenic organism according to the present invention. In particular, the term "homologue" covers homology with respect to structure and/or function providing the resultant nucleotide sequence has the ability to act as or code for a terminator or signal respectively. With respect to sequence homology, preferably there is at least 75%, more preferably at least 85%, more preferably at least 90% homology to SEQ ID NO.s 13 and 14 (respectively) shown in the attached sequence listings. More preferably there is at least 95%, more preferably at least 98%, homology to SEQ ID NO.s 13 and 14 (respectively) shown in the attached sequence listings.

The terms "variant", "homologue" or "fragment" in relation to the signal amino acid sequence include any substitution of, variation of, modification of, replacement of, deletion of or addition of one (or more) amino acid from or to the sequence providing the resultant sequence has the ability to act as a signal sequence in an expression system - such as the transformed cell or the transgenic organism according to the present invention. In particular, the term "homologue" covers homology with respect to structure and/or function providing the resultant nucleotide sequence has the ability to act as or code for a signal respectively. With respect to sequence homology, preferably there is at least 75%, more preferably at least 85%, more preferably at least 90% homology to SEQ ID NO 15 shown in the attached sequence listings. More preferably there is at least

95%, more preferably at least 98%, homology to SEQ ID NO 15 shown in the attached sequence listings.

The above terms are synonymous with allelic variations of the sequences.

5

The term "complementary" means that the present invention also covers nucleotide sequences that can hybridise to the nucleotide sequences of the coding sequence or the promoter sequence, respectively.

10 The term "nucleotide" in relation to the present invention includes genomic DNA, cDNA, synthetic DNA, and RNA. Preferably it means DNA, more preferably cDNA for the coding sequence of the present invention.

15 The term "construct" - which is synonymous with terms such as "conjugate", "cassette" and "hybrid" - includes a GOI directly or indirectly attached to a promoter. An example of an indirect attachment is the provision of a suitable spacer group such as an intron sequence, such as the *Sh1*-intron or the ADH intron, intermediate the promoter and the GOI. The same is true for the term "fused" in relation to the present invention which includes direct or indirect attachment. In each case, it is highly preferred that the terms
20 do not cover the natural combination of the gene coding for the enzyme ordinarily associated with the wild type gene promoter and when they are both in their natural environment. A highly preferred embodiment is the or a GOI being operatively linked to a or the promoter.

25 The construct may even contain or express a marker which allows for the selection of the genetic construct in, for example, a filamentous fungus, preferably of the genus *Aspergillus*, such as *Aspergillus niger*, or plants, preferably cereals, such as maize, rice, barley etc., into which it has been transferred. Various markers exist which may be used, such as for example those encoding mannose-6-phosphate isomerase (especially for
30 plants) or those markers that provide for antibiotic resistance - e.g. resistance to G418, hygromycin, bleomycin, kanamycin and gentamycin.

The term "vector" includes expression vectors and transformation vectors.

The term "expression vector" means a construct capable of *in vivo* or *in vitro* expression.

- 5 The term "transformation vector" means a construct capable of being transferred from one species to another - such as from an *E.coli* plasmid to a filamentous fungus, preferably of the genus *Aspergillus*. It may even be a construct capable of being transferred from an *E.coli* plasmid to an *Agrobacterium* to a plant.
- 10 The term "tissue" includes tissue *per se* and organ.

- The term "organism" in relation to the present invention includes any organism that could comprise the promoter according to the present invention and/or the nucleotide sequence coding for the enzyme according to the present invention and/or products obtained therefrom, wherein the promoter can allow expression of a GOI and/or wherein the nucleotide sequence according to the present invention can be expressed when present in the organism.

- 20 Preferably the organism is a filamentous fungus, preferably of the genus *Aspergillus*, more preferably *Aspergillus niger*.

- The term "transgenic organism" in relation to the present invention includes any organism that comprises the promoter according to the present invention and/or the nucleotide sequence coding for the enzyme according to the present invention and/or products obtained therefrom, wherein the promoter can allow expression of a GOI and/or wherein the nucleotide sequence according to the present invention can be expressed within the organism. Preferably the promoter and/or the nucleotide sequence is (are) incorporated in the genome of the organism.

- 30 Preferably the transgenic organism is a filamentous fungus, preferably of the genus *Aspergillus*, more preferably *Aspergillus niger*.

~~TECH CENTER 1600/2900~~

Therefore, the transgenic organism of the present invention includes an organism comprising any one of, or combinations of, the promoter according to the present invention, the nucleotide sequence coding for the enzyme according to the present invention, constructs according to the present invention, vectors according to the present invention, plasmids according to the present invention, cells according to the present invention, tissues according to the present invention or the products thereof. For example the transgenic organism can comprise a GOI, preferably an exogenous nucleotide sequence, under the control of the promoter according to the present invention. The transgenic organism can also comprise the nucleotide sequence coding for the enzyme of the present invention under the control of a promoter, which may be the promoter according to the present invention.

In a highly preferred embodiment, the transgenic organism does not comprise the combination of the promoter according to the present invention and the nucleotide sequence coding for the enzyme according to the present invention, wherein both the promoter and the nucleotide sequence are native to that organism and are in their natural environment. Thus, in these highly preferred embodiments, the present invention does not cover the native nucleotide coding sequence according to the present invention in its natural environment when it is under the control of its native promoter which is also in its natural environment. In addition, in this highly preferred embodiment, the present invention does not cover the native enzyme according to the present invention when it is in its natural environment and when it has been expressed by its native nucleotide coding sequence which is also in its natural environment and when that nucleotide sequence is under the control of its native promoter which is also in its natural environment.

25

The term "promoter" is used in the normal sense of the art, e.g. an RNA polymerase binding site in the Jacob-Mond theory of gene expression.

In one aspect, the promoter of the present invention is capable of expressing a GOI, 30 which can be the nucleotide sequence coding for the enzyme of the present invention.

In another aspect, the nucleotide sequence according to the present invention is under the control of a promoter that allows expression of the nucleotide sequence. In this regard, the promoter need not necessarily be the same promoter as that of the present invention. In this aspect, the promoter may be a cell or tissue specific promoter. If, for example, 5 the organism is a plant then the promoter can be one that affects expression of the nucleotide sequence in any one or more of stem, sprout, root and leaf tissues.

By way of example, the promoter for the nucleotide sequence of the present invention can be the α -Amy 1 promoter (otherwise known as the Amy 1 promoter, the Amy 637 10 promoter or the α -Amy 637 promoter) as described in our co-pending UK patent application No. 9421292.5 filed 21 October 1994. That promoter comprises the sequence shown in Figure 1.

Alternatively, the promoter for the nucleotide sequence of the present invention can be 15 the α -Amy 3 promoter (otherwise known as the Amy 3 promoter, the Amy 351 promoter or the α -Amy 351 promoter) as described in our co-pending UK patent application No. 9421286.7 filed 21 October 1994. That promoter comprises the sequence shown in Figure 2.

20 Preferably, the promoter is the promoter of the present invention.

In addition to the nucleotide sequences described above, the promoters, particularly that of the present invention, could additionally include features to ensure or to increase expression in a suitable host. For example, the features can be conserved regions such 25 as a Pribnow Box or a TATA box. The promoters may even contain other sequences to affect (such as to maintain, enhance, decrease) the levels of expression of the GOI. For example, suitable other sequences include the *Sh1*-intron or an ADH intron. Other sequences include inducible elements - such as temperature, chemical, light or stress inducible elements.

30

Also, suitable elements to enhance transcription or translation may be present. An example of the latter element is the TMV 5' signal sequence (see Sleat Gene 217 [1987]

217-225; and Dawson Plant Mol. Biol. 23 [1993] 97.

In addition the present invention also encompasses combinations of promoters and/or nucleotide sequences coding for proteins or enzymes and/or elements. For example, the
5 present invention encompasses the combination of a promoter according to the present invention operatively linked to a GOI, which could be a nucleotide sequence according to the present invention, and another promoter such as a tissue specific promoter operatively linked to the same or a different GOI.

10 The present invention also encompasses the use of promoters to express a nucleotide sequence coding for the enzyme according to the present invention, wherein a part of the promoter is inactivated but wherein the promoter can still function as a promoter. Partial inactivation of a promoter in some instances is advantageous.

15 In particular, with the Amy 351 promoter mentioned earlier it is possible to inactivate a part of it so that the partially inactivated promoter expresses GOIs in a more specific manner such as in just one specific tissue type or organ.

20 The term "inactivated" means partial inactivation in the sense that the expression pattern of the promoter is modified but wherein the partially inactivated promoter still functions as a promoter. However, as mentioned above, the modified promoter is capable of expressing a GOI in at least one (but not all) specific tissue of the original promoter. One such promoter is the Amy 351 promoter described above.

25 Examples of partial inactivation include altering the folding pattern of the promoter sequence, or binding species to parts of the nucleotide sequence, so that a part of the nucleotide sequence is not recognised by, for example, RNA polymerase. Another, and preferable, way of partially inactivating the promoter is to truncate it to form fragments thereof. Another way would be to mutate at least a part of the sequence so that the RNA
30 polymerase can not bind to that part or another part.

Another modification is to mutate the binding sites for regulatory proteins for example the CreA protein known from filamentous fungi to exert carbon catabolite repression, and thus abolish the catabolite repression of the native promoter.

- 5 The term "GOI" with reference to the present invention means any gene of interest. A GOI can be any nucleotide that is either foreign or natural to the organism (e.g. filamentous fungus, preferably of the genus *Aspergillus*, or a plant) in question. Typical examples of a GOI include genes encoding for proteins and enzymes that modify metabolic and catabolic processes. The GOI may code for an agent for introducing or
10 increasing pathogen resistance. The GOI may even be an antisense construct for modifying the expression of natural transcripts present in the relevant tissues. The GOI may even code for a non-natural protein of a filamentous fungus, preferably of the genus *Aspergillus*, or a compound that is of benefit to animals or humans.
- 15 For example, the GOI could code for a pharmaceutically active protein or enzyme such as any one of the therapeutic compounds insulin, interferon, human serum albumin, human growth factor and blood clotting factors. In this regard, the transformed cell or organism could prepare acceptable quantities of the desired compound which would be easily retrievable from, the cell or organism. The GOI may even be a protein giving
20 nutritional value to a food or crop. Typical examples include plant proteins that can inhibit the formation of anti-nutritive factors and plant proteins that have a more desirable amino acid composition (e.g. a higher lysine content than a non-transgenic plant). The GOI may even code for an enzyme that can be used in food processing such as chymosin, thaumatin and α -galactosidase. The GOI can be a gene encoding for any one of a pest
25 toxin, an antisense transcript such as that for patatin or α -amylase, ADP-glucose pyrophosphorylase (e.g. see EP-A-0455316), a protease antisense or a glucanase.

The GOI can be the nucleotide sequence coding for the α -amylase enzyme which is the subject of our co-pending UK patent application 9413439.2 filed on 4 July 1994, the
30 sequence of which is shown in Figure 3. The GOI can be the nucleotide sequence coding for the α -amylase enzyme which is the subject of our co-pending UK patent application 9421290.9 filed on 21 October 1994, the sequence of which is shown in Figure 4. The

GOI can be any of the nucleotide sequences coding for the ADP-glucose pyrophosphorylase enzymes which are the subject of our co-pending PCT patent application PCT/EP94/01082 filed 7 April 1994, the sequences of which are shown in Figures 5 and 6. The GOI can be any of the nucleotide sequences coding for the α -glucan lyase enzyme which are described in our co-pending PCT patent application PCT/EP94/03397 filed 15 October 1994, the sequences of which are shown in Figures 7-10.

In one preferred embodiment, the GOI is a nucleotide sequence coding for the enzyme according to the present invention.

As mentioned above, a preferred host organism is of the genus *Aspergillus*, such as *Aspergillus niger*. The transgenic *Aspergillus* according to the present invention can be prepared by following the teachings of Rambosek,J. and Leach,J. 1987 (Recombinant DNA in filamentous fungi: Progress and Prospects. CRC Crit. Rev. Biotechnol. 6:357-393), Davis R.W. 1994 (Heterologous gene expression and protein secretion in *Aspergillus*. In: Martinelli S.D., Kinghorn J.R.(Editors) *Aspergillus*: 50 years on. Progress in industrial microbiology vol 29. Elsevier Amsterdam 1994. pp 525-560), Ballance,D.J. 1991 (Transformation systems for Filamentous Fungi and an Overview of Fungal Gene structure. In :Leong,S.A., Berka R.M. (Editors) Molecular Industrial Mycology. Systems and Applications for Filamentous Fungi. Marcel Dekker Inc. New York 1991. pp 1-29) and Turner G. 1994 (Vectors for genetic manipulation. In: Martinelli S.D., Kinghorn J.R.(Editors) *Aspergillus*: 50 years on. Progress in industrial microbiology vol 29. Elsevier Amsterdam 1994. pp. 641-666). However, the following commentary provides a summary of those teachings for producing transgenic *Aspergillus* according to the present invention.

Filamentous fungi have during almost a century been widely used in industry for production of organic compounds and enzymes. Traditional Japanese koji and soy fermentations have used *Aspergillus* sp for hundreds of years. In this century *Aspergillus niger* has been used for production of organic acids particular citric acid and for production of various enzymes for use in industry.

There are two major reasons for that filamentous fungi have been so widely used in industry. First filamentous fungi can produce high amounts of extracellular products, for example enzymes and organic compounds such as antibiotics or organic acids. Second filamentous fungi can grow on low cost substrates such as grains, bran, beet pulp etc.

- 5 The same reasons have made filamentous fungi attractive organisms as hosts for heterologous expression according to the present invention.

In order to prepare the transgenic *Aspergillus*, expression constructs are prepared by inserting a GOI (such as an amylase or SEQ. I.D. No. 2) into a construct designed for

- 10 expression in filamentous fungi.

Several types of constructs used for heterologous expression have been developed. The constructs contain the promoter according to the present invention (or if desired another promoter if the GOI codes for the enzyme according to the present invention) which is

- 15 active in fungi. Examples of promoters other than that of the present invention include a fungal promoter for a highly expressed extracellular enzyme, such as the glucoamylase promoter or the α -amylase promoter. The GOI can be fused to a signal sequence (such as that of the present invention or another suitable sequence) which directs the protein encoded by the GOI to be secreted. Usually a signal sequence of fungal origin is used,

- 20 such as that of the present invention. A terminator active in fungi ends the expression system, such as that of the present invention.

Another type of expression system has been developed in fungi where the GOI is fused to a smaller or a larger part of a fungal gene encoding a stable protein. This can stabilize

- 25 the protein encoded by the GOI. In such a system a cleavage site, recognized by a specific protease, can be introduced between the fungal protein and the protein encoded by the GOI, so the produced fusion protein can be cleaved at this position by the specific protease thus liberating the protein encoded by the GOI ("POI"). By way of example, one can introduce a site which is recognized by a KEX-2 like peptidase found in at least
30 some *Aspergilli*. Such a fusion leads to cleavage *in vivo* resulting in protection of the POI and production of POI and not a larger fusion protein.

Heterologous expression in *Aspergillus* has been reported for several genes coding for bacterial, fungal, vertebrate and plant proteins. The proteins can be deposited intracellularly if the GOI is not fused to a signal sequence. Such proteins will accumulate in the cytoplasm and will usually not be glycosylated which can be an advantage for some 5 bacterial proteins. If the GOI is equipped with a signal sequence the protein will accumulate extracellularly.

With regard to product stability and host strain modifications, some heterologous proteins are not very stable when they are secreted into the culture fluid of fungi. Most fungi 10 produce several extracellular proteases which degrade heterologous proteins. To avoid this problem special fungal strains with reduced protease production have been used as host for heterologous production.

For the transformation of filamentous fungi, several transformation protocols have been 15 developed for many filamentous fungi (Ballance 1991, *ibid*). Many of them are based on preparation of protoplasts and introduction of DNA into the protoplasts using PEG and Ca²⁺ ions. The transformed protoplasts then regenerate and the transformed fungi are selected using various selective markers. Among the markers used for transformation are a number of auxotrophic markers such as *argB*, *trpC*, *niaD* and *pyrG*, antibiotic 20 resistance markers such as benomyl resistance, hygromycin resistance and phleomycin resistance. A very common used transformation marker is the *amdS* gene of *A. nidulans* which in high copy number allows the fungus to grow with acrylamide as the sole nitrogen source.

25 Even though the enzyme, the nucleotide sequence coding for same and the promoter of the present invention are not disclosed in EP-B-0470145 and CA-A-2006454, those two documents do provide some useful background commentary on the types of techniques that may be employed to prepare transgenic plants according to the present invention. Some of these background teachings are now included in the following commentary.

30

The basic principle in the construction of genetically modified plants is to insert genetic information in the plant genome so as to obtain a stable maintenance of the inserted

genetic material.

Several techniques exist for inserting the genetic information, the two main principles being direct introduction of the genetic information and introduction of the genetic
5 information by use of a vector system. A review of the general techniques may be found in articles by Potrykus (*Annu Rev Plant Physiol Plant Mol Biol* [1991] 42:205-225) and Christou (Agro-Food-Industry Hi-Tech March/April 1994 17-27).

Thus, in one aspect, the present invention relates to a vector system which carries a
10 promoter or nucleotide sequence or construct according to the present invention and which is capable of introducing the promoter or nucleotide sequence or construct into the genome of an organism, such as a plant.

The vector system may comprise one vector, but it can comprise two vectors. In the case
15 of two vectors, the vector system is normally referred to as a binary vector system. Binary vector systems are described in further detail in Gynheung An et al. (1980), *Binary Vectors, Plant Molecular Biology Manual A3*, 1-19.

One extensively employed system for transformation of plant cells with a given promoter
20 or nucleotide sequence or construct is based on the use of a Ti plasmid from *Agrobacterium tumefaciens* or a Ri plasmid from *Agrobacterium rhizogenes* An et al. (1986). *Plant Physiol.* 81, 301-305 and Butcher D.N. et al. (1980). *Tissue Culture Methods for Plant Pathologists*, eds.: D.S. Ingrams and J.P. Helgeson, 203-208.

25 Several different Ti and Ri plasmids have been constructed which are suitable for the construction of the plant or plant cell constructs described above. A non-limiting example of such a Ti plasmid is pGV3850.

30 The promoter or nucleotide sequence or construct of the present invention should preferably be inserted into the Ti-plasmid between the terminal sequences of the T-DNA or adjacent a T-DNA sequence so as to avoid disruption of the sequences immediately surrounding the T-DNA borders, as at least one of these regions appear to be essential

for insertion of modified T-DNA into the plant genome.

As will be understood from the above explanation, if the organism is a plant, then the vector system of the present invention is preferably one which contains the sequences
5 necessary to infect the plant (e.g. the *vir* region) and at least one border part of a T-DNA sequence, the border part being located on the same vector as the genetic construct.

Furthermore, the vector system is preferably an *Agrobacterium tumefaciens* Ti-plasmid or an *Agrobacterium rhizogenes* Ri-plasmid or a derivative thereof, as these plasmids are
10 well-known and widely employed in the construction of transgenic plants, many vector systems exist which are based on these plasmids or derivatives thereof.

In the construction of a transgenic plant the promoter or nucleotide sequence or construct of the present invention may be first constructed in a microorganism in which the vector
15 can replicate and which is easy to manipulate before insertion into the plant. An example of a useful microorganism is *E. coli*, but other microorganisms having the above properties may be used. When a vector of a vector system as defined above has been constructed in *E. coli*, it is transferred, if necessary, into a suitable *Agrobacterium* strain, e.g. *Agrobacterium tumefaciens*. The Ti-plasmid harbouring the promoter or nucleotide sequence or construct of the invention is thus preferably transferred into a suitable *Agrobacterium* strain, e.g. *A. tumefaciens*, so as to obtain an *Agrobacterium* cell harbouring the promoter or nucleotide sequence or construct of the invention, which DNA is subsequently transferred into the plant cell to be modified.
20

25 As reported in CA-A-2006454, a large amount of cloning vectors are available which contain a replication system in *E. coli* and a marker which allows a selection of the transformed cells. The vectors contain for example pBR 322, pUC series, M13 mp series, pACYC 184 etc.

30 In this way, the nucleotide or construct or promoter of the present invention can be introduced into a suitable restriction position in the vector. The contained plasmid is used for the transformation in *E. coli*. The *E. coli* cells are cultivated in a suitable nutrient

medium and then harvested and lysed. The plasmid is then recovered. As a method of analysis there is generally used sequence analysis, restriction analysis, electrophoresis and further biochemical-molecular biological methods. After each manipulation, the used DNA sequence can be restricted and connected with the next DNA sequence. Each 5 sequence can be cloned in the same or different plasmid.

After each introduction method of the desired promoter or construct or nucleotide sequence according to the present invention in the plants the presence and/or insertion of further DNA sequences may be necessary. If, for example, for the transformation the 10 Ti- or Ri-plasmid of the plant cells is used, at least the right boundary and often however the right and the left boundary of the Ti- and Ri-plasmid T-DNA, as flanking areas of the introduced genes, can be connected. The use of T-DNA for the transformation of plant cells has been intensively studied and is described in EP-A-120516; Hoekema, in: The Binary Plant Vector System Offset-drukkerij Kanters B.B., Albllasserdam, 1985, 15 Chapter V; Fraley, et al., Crit. Rev. Plant Sci., 4:1-46; and An et al., EMBO J. (1985) 4:277-284.

Direct infection of plant tissues by *Agrobacterium* is a simple technique which has been widely employed and which is described in Butcher D.N. et al. (1980), *Tissue Culture* 20 *Methods for Plant Pathologists*, eds.: D.S. Ingrams and J.P. Helgeson, 203-208. For further teachings on this topic see Potrykus (Annu Rev Plant Physiol Plant Mol Biol [1991] 42:205-225) and Christou (Agro-Food-Industry Hi-Tech March/April 1994 17-27). With this technique, infection of a plant may be done on a certain part or tissue of the plant, i.e. on a part of a leaf, a root, a stem or another part of the plant.

25

Typically, with direct infection of plant tissues by *Agrobacterium* carrying the promoter and/or the GOI, a plant to be infected is wounded, e.g. by cutting the plant with a razor or puncturing the plant with a needle or rubbing the plant with an abrasive. The wound is then inoculated with the *Agrobacterium*. The inoculated plant or plant part is then 30 grown on a suitable culture medium and allowed to develop into mature plants.

When plant cells are constructed, these cells may be grown and maintained in accordance with well-known tissue culturing methods such as by culturing the cells in a suitable culture medium supplied with the necessary growth factors such as amino acids, plant hormones, vitamins, etc.

5

Regeneration of the transformed cells into genetically modified plants may be accomplished using known methods for the regeneration of plants from cell or tissue cultures, for example by selecting transformed shoots using an antibiotic and by subculturing the shoots on a medium containing the appropriate nutrients, plant hormones, etc.

10

Further teachings on plant transformation may be found in EP-A-0449375.

In summation, the present invention provides an arabinofuranosidase enzyme having 15 arabinoxylan degrading activity and the nucleotide sequence coding for the same. In addition, it provides a promoter that can control the expression of that, or another, nucleotide sequence. In addition it includes terminator and signal sequences for the same.

20

The following sample was deposited in accordance with the Budapest Treaty at the recognised depositary The National Collections of Industrial and Marine Bacteria Limited (NCIMB) at 23 St. Machar Drive, Aberdeen, Scotland, United Kingdom. AB2 1RY on 16 January 1995:

25

E.coli containing plasmid pB53.1 {i.e. *E.coli* DH5 α -pB53.1}. The deposit number is NCIMB 40703.

The present invention will now be described by way of example.

30

In the following Examples reference is made to the accompanying figures in which:

Figures 1-10 are sequences of promoters and GOIs of earlier patent applications that are useful for use with the aspects of the present invention;

5 Figure 11 is a plasmid map of the plasmid pBS3.1, which is the subject of deposit NCIMB 40703;

Figure 12 is a schematic diagram of deletions made to the promoter of the present invention;

10 Figure 13 is a plasmid map of pXP-AMY;

Figure 14 is a plasmid map of pXP-XssAMY;

15 Figure 15 is a graph;

Figure 16 is an HP-TLC profile;

Figure 17 is an HP-TLC profile;

20 Figure 18 is an HPLC profile;

Figure 19 is a viscosity plot;

Figure 20 is an activity plot;

25

Figure 21 is an activity plot; and

Figure 22 is an activity plot.

30 The following text discusses the use of *inter alia* recombinant DNA techniques. General teachings of recombinant DNA techniques may be found in Sambrook, J., Fritsch, E.F., Maniatis T. (Editors) Molecular Cloning. A laboratory manual. Second edition. Cold

Spring Harbour Laboratory Press, New York 1989.

In these Examples, the enzyme of the present invention is sometimes referred to as AbfC.

In addition, the promoter of the present invention is sometimes referred to as the AbfC

5 promoter.

Purification of the arabinofuranosidase

Aspergillus niger 3M43 was grown in medium containing wheat bran and beet pulp. The
10 fermentation broth was separated from the solid part of the broth by filtration. Concentrated fermentation broth was loaded on a 25X100mm Q-SEPHAROSE (Pharmacia) high Performance column, equilibrated with 20 mM Tris, HCl pH 7.5, and a linear gradient from 0-500 Mm NaCl was performed and fractions of the eluate was collected. The Arabinofuranosidase was eluted at 130-150 Mm NaCl.

15

The fractions containing the arabinofuranosidase were combined and desalting using a 50x200 mm G-25 SEPHAROSE Superfine (Pharmacia). The column was eluted with distilled water.

20

After desalting the enzyme was concentrated using High-Trap spin columns. Next the concentrated and desalting fractions were subjected to gel filtration on a 50x600 mm SUPERDEX 50 column. The sample was loaded and the column was eluted with 0.2 M Phosphate buffer pH 7.0 plus 0.2 M NaCl, and fractions of the eluate were collected.

25

The fractions containing arabinofuranosidase were combined and desalting and concentrated as described above. The combined fractions were loaded on a 16X100 mm Phenylsepharose High Performance column (Pharmacia), equilibrated with 50 mM Phosphate buffer pH 6.0, containing 1.5 M $(\text{NH}_4)_2\text{SO}_4$. A gradient where the $(\text{NH}_4)_2\text{SO}_4$ concentration was varied from 1.5 - 0 M was applied and the eluate collected in fractions.

30

The fractions containing Arabinofuranosidase were combined. The purity of the arabinofuranosidase was evaluated by SDS-PAGE using the Phast system gel (Pharmacia).

Characterization

The molecular weight of the purified arabinofuranosidase was determined by mass spectrometry using laser desorption technology. The MW of the arabinofuranosidase was
5 found to be 33.270 D ± 50 D.

The pI value was determined by use of a Broad pI Kit (Pharmacia). The arabinofuranosidase has a pI value of about 3.7.

10 After SDS-PAGE analysis, treatment PAS reagent showed that the arabinofuranosidase was glycosylated. The PAS staining was done according to the procedure of I. Van-Seuningen and M. Davril (1992) Electrophoresis 13 pp 97-99.

Activity Studies

15 Activity of AbfC as a function of water soluble pentosan (WSP) concentrations (mg/ml) was determined. The results are shown in Figure 21. The results show that AbfC activity reached maximum at substrate concentration of 8 mg/ml WSP.

pH Activity Studies

20 The effect of pH on the activity of the arabinofuranosidase of the present invention was investigated using water soluble pentosan (10 mg/ml) from wheat as a substrate in 50 mM citric acid sodium phosphate buffer. The incubation time was 15 minutes. The arabinofuranosidase of the present invention was observed to have a wide pH optima range of from about 2.5 to about 7.0 (see Figure 20), more especially from about 3.3 to
25 about 4.6, more especially about 4.

Temperature Activity Studies

30 The effect of temperature on the activity of the arabinofuranosidase of the present invention was investigated using water soluble pentosan (10 mg/ml) from wheat as a substrate in 50 mM sodium acetate at a pH of 5.0. The incubation time was 15 minutes.

The arabinofuranosidase of the present invention was observed to have an optimal activity at a temperature of from about 40°C to about 60°C, more especially from about 45°C to about 55°C, more especially about 50°C (Figure 22). The enzyme is still active at about 10°C and showed residual activity at 70°C and 80°C.

5

Amino acid sequencing of the arabinofuranosidase

The enzyme was digested with endoproteinase Lys-C sequencing grade from Boehringer Mannheim using a modification of the method described by Stone & Williams 1993
10 (Stone, K.L. and Williams, K.R. (1993). Enzymatic digestion of Proteins and HPLC Peptide Isolation. In : Matsudaira P. (Editor). A practical Guide to Protein and Peptide Purification for Microsequencing. Second Edition. Academic Press, San Diego 1993. pp 45-73).

15 Freeze dried β -arabinofuranosidase (0.4 mg) was dissolved in 50 μ l of 8M urea, 0.4 M NH₄HCO₃, pH 8.4. After overlay with N₂ and addition of 5 μ l of 45 Mm DTT, the protein was denatured and reduced for 15 min at 50°C under N₂. After cooling to RT, 5 μ l of 100 Mm iodoacetamide was added for the cysteines to be derivatised for 15 min at RT in the dark under N₂. Subsequently, 90 μ l of water and 5 μ g of endoproteinase
20 Lys-C in 50 μ l of 50 Mm Tricine and 10 mM EDTA, pH 8.0, was added and the digestion was carried out for 24h at 37°C under N₂. The resulting peptides were separated by reversed phase HPLC on a VYDAC C18 column (0.46 x 15 cm; 10 μ m; The Separations Group; California) using solvent A: 0.1% TFA in water and solvent B: 0.1% TFA in acetonitrile. Selected peptides were rechromatographed on a Develosil C18
25 column (0.46 x 10 cm; 3 μ m) using the same solvent system prior to sequencing on an Applied Biosystems 476A sequencer using pulsed-liquid fast cycles.

The following peptide sequences were found:

- 30 SEQ I.D. No. 4
SEQ I.D. No. 5
SEQ I.D. No. 6

SEQ I.D. No. 7

SEQ I.D. No. 8

Isolation of a PCR clone of a fragment of the gene

5

PCR primers were synthesised using an Applied Biosystems DNA synthesiser model 392. In this regard, PCR primers were synthesized from one of the found peptide sequences, namely SEQ ID No. 5. The primers were:

10 One primer from EMTAQQA (reversed)

SEQ ID NO. 9 GCY TGN GCN GTC ATY TC
 17 mer 64 mix

15 One primer from MIVEAIG

SEQ ID NO. 10 ATG ATH GTN GAR GCN ATH GG
 20 mer 288 mix

20 PCR amplification was performed with 100 pmol of each of these primers in 100 µl reactions using AmpliTaq polymerase (PERKIN ELMER). The following program was:

	<u>STEP</u>	<u>TEMP</u>	<u>TIME</u>
25	1	94°C	2 min
	2	94°C	1 min
	3	55°C	2 min
	4	72°C	2 min
	5	72°C	5 min
30	6	5°C	SOAK

Steps 2-4 were repeated for 40 cycles.

PCR reactions were performed on a PERKIN ELMER DNA Thermal Cycler.

A 100 bp amplified fragment was isolated and cloned into a pT7-Blue T-vector, according to the manufacturers instructions (Novagen).

5

Isolation of *A. niger* genomic DNA

1g. of frozen *A. niger* mycelium was ground in a mortar under liquid nitrogen. Following evaporation of the nitrogen cover, the ground mycelium was extracted with 15ml of
10 an extraction buffer (100mM Tris-HCl, pH 8.0, 0.50mM EDTA, 500mM NaCl, 10mM β-mercaptoethanol) containing 1ml 20% sodium dodecyl sulphate. After incubation at 65°C for 10 min. 5ml 5M KAc, pH 5.0, was added and the mixture further incubated, after mixing, on ice for 20 mins. After extraction, the mixture was centrifuged for 20 mins. and the supernatant mixed with 0.6 vol. isopropanol to precipitate the extracted
15 DNA. After further centrifugation for 15 mins. the DNA pellet was dissolved in 0.7 ml TE (10mM Tris, HCl pH 8.0, 1mM EDTA) and precipitated with 75 µl 3M NaAc, pH 4.8, and 500 µl isopropanol.

After centrifugation the pellet was washed with 70% ETOH and dried under vacuum.
20 The DNA was dissolved in 200 µl TE and stored at -20°C.

Construction of a library

20 µg genomic DNA was partly digested with Tsp509I, which gives ends which are
25 compatible with *EcoRI* ends. The digested DNA was separated on a 1 % agarose gel and fragments of 4-10 kb was purified. A λZAPII *EcoRI*/CIAP kit from Stratagene was used for library construction according to the manufacturers instructions. 2 µl of the ligation (totally 5 µl) was packed with Gigapack Gold II packing extract from Stratagene. The library contained 650,000 independent clones.

30

Screening of the library

2 X 50.000 pfu was plated on NZY plates and plaquelifts were done on Hybond N sheets (Amersham). Plaquelifts were done in duplicates. The sheets were hybridized with the 5 PCR clone labelled with ^{32}P dCTP (Amersham) using Ready-to-go labelling kit from Pharmacia. Positive clones were reckoned only when hybridization was detected on both sheets. The gene was sequenced, and the found sequence showed that all of the peptides sequenced were coded by the found sequence.

10 Sequence information

SEQ. ID. No. 12 presents the promoter sequence, the enzyme coding sequence, the terminator sequence and the signal sequence and the amino acid sequence of the enzyme of the present invention.

15

Arabinofuranosidase assay

Two different arabinoxylan preparations from wheat flour, Wheat Insoluble Pentosan (WIP) and Wheat soluble Pentosan (WSP), were degraded with the arabinofuranosidase 20 enzyme of the present invention alone and in combination with an endoxylanase purified from *A. niger*. The assays were done on 1% substrate in 50 Mm 50 Mm Na-acetate buffer at pH 5.0. The reactions were performed at 30 °C for 2.5 hours. The reactions were stopped by addition of 3 vol. ethanol which precipitates the high molecular weight material. The samples were centrifuged and the supernatants were collected, dried under 25 vacuum and resuspended in 0.5 ml distilled water. The samples were diluted 1:1 in water and analysed on a Chromopack Carbohydrate Pb column (300X7.8 mm, cat. 29010) using Shimadzu C-R4A Chromatopac HPLC system using a Shimadzu RI D-6A refractive index detector in accordance with the suppliers instructions.

30 The column was calibrated using a standard composed of 0.48 mg/ml xylotriose, 0.48 mg/ml xylobiose, 0.60 mg/ml xylose and 0.58 mg/ml L-arabinose. The peaks were identified and quantified using the software supplied with the equipment.

Results - Liberated saccharides from Wheat Insoluble Pentosan

Substrate 1% WIP in 50 Mm Na-acetate buffer pH 5.0. Values are expressed in mg/ml.

	xylotriose	xylobiose	xylose	arabinose
no enzyme	0.0	0.0	0.0	0.0
abfC	0.0	0.0	0.0	0.11
xyl	0.09	0.14	0.0	0.0
abfC + xyl	0.37	0.41	0.0	0.30

10

abfC denotes the enzyme according to the present invention; and xyl denotes the xylanase described before.

Results - Saccharides liberated from Wheat Soluble Pentosan

15

Substrate 1 % WSP in 50 Mm Na-acetate buffer pH 5.0. Values are expressed in mg/ml.

	xylotriose	xylobiose	xylose	arabinose
no enzyme	0.0	0.0	0.0	0.0
abfC	0.0	0.0	0.0	0.30
xyl	0.08	0.14	0.0	0.0
abfC + xyl	0.42	0.47	0.0	0.42

25

abfC denotes the enzyme according to the present invention; and xyl denotes the xylanase described before.

Figure 17 shows HP-TLC profiles of the AbfC enzyme acting synergistically with Xylanase A. In this Figure, the following abbreviations are used: water-soluble pentosan (WSP); water-insoluble pentosan (WIP); and oat xylan as substrate. The standards were: X- xylose; X₂- xylobiose; X₃- xylotriose; A- arabinose.

5

Figure 18 shows the HPLC analysis of hydrolysis products using 1% oat spelt xylan as the substrate. Figure 18(a) and Figure 18(b) show the products when the AbfC enzyme and the xylanase enzyme respectively were used alone. Figure 18(c) show the products when the AbfC enzyme and the xylanase enzyme when combined.

10

The results of these experiments provide two important findings.

First the enzyme of the present invention liberates arabinose, in particular L-arabinose, from arabinoxylan.

15

Second the combined actions of the enzyme according to the present invention with the endoxylanase is significantly higher than the sum of their individual action. Accordingly, the two enzymes affect each others enzymatic activities in a synergistic fashion.

20 Induction of the AbfC gene: Identification of inducers

The regulation of transcription of the AbfC encoding gene of *Aspergillus niger* was studied using a strain containing a fusion of the AbfC promoter to the β -glucuronidase encoding gene (uid A) of *E. coli*.

25

GUS producing transformants were grown on different carbon sources and assayed both qualitatively and quantitatively for the ability to hydrolyse p-nitrophenol glucuronide.

The results are shown below:

30

CARBON SOURCE GUS ACTIVITY AFTER 24 HOURS INDUCTION

(1%) (units/mg)

	xylose	12.37
5	xylitol	1.49
	arabinose	6.66
	arabitol	5.30
	glucose	0.70
	cellubiose	0.95
10	xylo-oligomer 70	17.26
	glucopyranoside	0.40
	methyl-xylopyranoside	24.20
	xyloglucan	1.00
	pectin	0.27
15	arabinogalactan	2.60
	arabitol + glucose	2.20

The results show that the AbfC promoter is switched on after 24 hours when grown in the presence of xylose, xylo-oligomer 70, methyl-xylopyranoside, arabinose and arabitol.

20 These studies also suggest that methyl-xylopyranoside is the natural and strongest inducer of this promoter.

The AbfC promoter is strongly repressed by glucose and is therefore under carbon catabolite repression. However, unlike all the published promoters for arabinofuranosidases, which are induced by arabinose and arabitol, the AbfC promoter of the present invention is regulated strongly by the intermediates in xylose metabolism. Accordingly, the present invention also covers an arabinofuranosidase promoter wherein the promoter is inducible by an intermediate in xylose metabolism.

Effects of different promoter deletions on the regulation of the expression of the AbfC gene

To study the regulation at the molecular level, experiments were set up to detect possible upstream regulating sequences required for expression of the AbfC gene. A series of plasmids with deletions in the 5' upstream region of the gene was constructed (see Figure 12). The *E coli* uid A gene was used as the reporter gene and a qualitative GUS assay was performed.

The results indicated that the truncated AbfC promoter of 590 bp contains sufficient information for the inducibility of the AbfC gene and its regulation. Deletion of 100 bps sequence from the *Xma* 111 to the *BamH*1 sites of the promoter led to a reduction in activity of this promoter. Therefore, this 100 bps area is important for good levels of gene expression. Deletion of 290 bps before the ATG identified this region to be important but not sufficient to abolish the activity of this promoter. All the transformants analysed containing this promoter construct showed very pale blue when tested (+- GUS). This region is as follows:

-170 T C A T C C A A T A T

As seen, this region contains the CCAAT element and is a putative target for a general transcriptional activator. This sequence is similar to the nuclear protein binding sites found in two starch inducible promoters: the *Aspergillus niger* glucoamylase gene and the *Aspergillus oryzae* amylase gene as well as the amdS gene of *Aspergillus nidulans*.

25

HETEROLOGOUS PROTEIN PRODUCTION USING *ASPERGILLUS NIGER*
TRANSFORMED WITH THE AbfC PROMOTER AND THE AbfC SIGNAL
SEQUENCE

5 Transformation of *Aspergillus Niger*

The protocol for transformation of *A. niger* was based on the teachings of Buxton, F.P., Gwynne D.I., Davis, R.W. 1985 (Transformation of *Aspergillus niger* using the *argB* gene of *Aspergillus nidulans*. Gene 37:207-214), Daboussi, M.J., Djeballi, A., Gerlinger, C.,
10 Blaiseau, P.L., Cassan, M., Lebrun, M.H., Parisot, D., Brygoo, Y. 1989 (Transformation of seven species of filamentous fungi using the nitrate reductase gene of *Aspergillus nidulans*. Curr. Genet. 15:453-456) and Punt, P.J., van den Hondel, C.A.M.J.J. 1992 (Transformation of filamentous fungi based on hygromycin B and Phleomycin resistance markers. Meth. Enzym. 216:447-457).

15

For the purification of protoplasts, spores from one PDA (Potato Dextrose Agar - from Difco Lab. Detroit) plate of fresh sporulated N400 (CBS 120.49, Centraalbureau voor Schimmelcultures, Baarn) (7 days old) are washed off in 5-10 ml water. A shake flask with 200 ml PDC (Potato Dextrose Broth, Difco 0549-17-9, Difco Lab. Detroit) is
20 inoculated with this spore suspension and shaken (250 rpm) for 16-20 hours at 30°C.

25 The mycelium is harvested using Miracloth paper and 3-4 g wet mycelium are transferred to a sterile petri dish with 10 ml STC (1.2 M sorbitol, 10 mM Tris HCl pH 7.5, 50 mM CaCl₂) with 75 mg lysing enzymes (Sigma L-2265) and 4500 units lyticase (Sigma L-8012).

30 The mycelium is incubated with the enzyme until the mycelium is degraded and the protoplasts are released. The degraded mycelium is then filtered through a sterile 60 µm mesh filter. The protoplasts are harvested by centrifugation 10 min at 2000 rpm in a swing out rotor. The supernatant is discarded and the pellet is dissolved in 8 ml 1.5 M MgSO₄, and then centrifuged at 3000 rpm for 10 min.

The upper band, containing the protoplasts is transferred to another tube, using a transfer pipette and 2 ml 0.6 M KCl is added. Carefully 5 ml 30% sucrose is added on the top and the tube is centrifuged 15 min at 3000 rpm.

- 5 The protoplasts, lying in the interface band, are transferred to a new tube and diluted with 1 vol. STC. The solution is centrifuged 10 min at 3000 rpm. The pellet is washed twice with STC, and finally solubilized in 1 ml STC. The protoplasts are counted and eventually concentrated before transformation.
- 10 For the transformation, 100 μ l protoplast solution (10^6 - 10^7 protoplasts) are mixed with 10 μ l DNA solution containing 5- 10 μ g DNA and incubated 25 min at room temperature. Then 60 % PEG-4000 is carefully added in portions of 200 μ l, 200 μ l and 800 μ l. The mixture is incubated 20 min at room temperature. 3 ml STC is added to the mixture and carefully mixed. The mixture is centrifuged 3000 rpm for 10 min.

15

The supernatant is removed and the protoplasts are solubilized in the remaining of the supernatant. 3-5 ml topagarose is added and the protoplasts are quickly spread on selective plates.

- 20 **AbfC promoter and heterologous gene expression**

The expression vector pXP-Amy (Figure 13) contains the 2.1 kb α -amylase encoding gene from *Thermomyces lanuginosus* cloned downstream of the AbfC promoter (2.1 kb) and upstream of the Xylanase A terminator. This vector together with the hygromycin gene as a selectable marker was used for co-transformation experiments to test the functionality of the AbfC promoter.

The best transformant was accumulated in shake flask experiments at least 1 gram per litre of α -amylase in the culture media. Starch degrading activity was then detected within 48 hours and a peak of enzyme activity is observed at 4 days of growth on sugar beet pulp and wheat bran (Figure 15).

AbfC signal sequence functions in protein secretion

An expression construct containing the signal peptide of the A bfC gene translationally fused to the mature α -amylase from *T. lanuginosus* was prepared and expression of this 5 construct in the production strains was observed. In this regard, the translational fusion construct pXPXss-Amy (Figure 14) was placed under the transcriptional control of the A bfC promoter and the xylanase A termination signal. The incorporation of an endogenous signal peptide resulted in increased detectability of co-transformants expressing both amylase and the hygromycin resistance marker. The endogenous signal 10 peptide directed the secretion of amylase out of the cell.

Substrate Specificity of A bfC Protein

The substrate specificity of the purified A bfC was determined using arabinose containing 15 hemicelluloses: arabinoxylans from wheat, oat and larch, branched and debranched arabinans; arabinogalactan, sugar beet pectin, and xyloglucan.

The HPLC and HP-TLC results are shown in Figure 16, in which the following abbreviations are used: WSP - water-soluble pentosan, WIP - water-insoluble pentosan, 20 AG - arabinogalactan, deB-A - debranched arabinan. The standards used were: A - arabinose, X - xylose.

The results indicate that arabinose is the hydrolysis product from arabinoxylans. No hydrolysis products were released from arabinogalactan, debranched arabinan or 25 xyloglucan. Arabinose was released as a hydrolysis product from branched arabinan. A bfC is therefore a 1,2/1,2 debranching enzyme and it has no activity towards linear 1,5 α -linked L-arabinofuranose residues found in debranched arabinans and arabinogalactan. This enzyme also releases a product when pectin is used as the substrate. It is believed that this product is an arabinose containing ferulic acid or an arabinobiose.

Reduction of Viscosity By AbfC

The results for the substrate specificity studies also suggest that the enzyme of the present invention could be used to reduce the viscosity of feeds. In this regard, the enzyme would reduce the viscosity of branched substrates by removing the branches but not the backbone of that substrate. This is in contrast to the known viscosity modifiers which degrade the substrate backbone.

Accordingly, the present invention covers a process of reducing the viscosity of a branched substrate wherein the enzyme degrades the branches of the substrate but not the backbone of the substrate.

In particular, the present invention covers the use of the enzyme of the present invention as a viscosity modifier.

In this regard, an experiment was carried out to investigate the reduction of viscosity of the water-soluble pentosan fraction from wheat flour by arabinofuranosidase. In this experiment, 6 ml water-soluble pentosan was incubated with 100 μ l of AbfC for 20 hours, 20°C at pH 5.5.

The results (see Figure 19) show that the enzyme of the present invention can be used to reduce the viscosity of pectins, especially pectins that are used in beverages - such as fruit juices.

Accordingly, the present invention covers the use of the enzyme of the present invention to reduce the viscosity of pectin.

ANTIBODY PRODUCTION

Antibodies were raised against the enzyme of the present invention by injecting rabbits with the purified enzyme and isolating the immunoglobulins from antiserum according to procedures described according to N Harboe and A Ingild ("Immunization, Isolation

of Immunoglobulins. Estimation of Antibody Titre" In A Manual of Quantitative Immunoelectrophoresis. Methods and Applications. N H Axelsen, *et al* (eds.). Universitetsforlaget, Oslo, 1973) and by T G Cooper ("The Tools of Biochemistry". John Wiley & Sons, New York, 1977).

5

SUMMARY

Even though it is known that *Aspergillus niger* produces arabinofuranosidases, the present invention provides a novel and inventive arabinofuranosidase, as well as the coding sequence therefor and the promoter for that sequence. An important advantage of the present invention is that the enzyme can be produced in high amounts.

In addition, the promoter and the regulatory sequences (such as the signal sequence and the terminator) can be used to express or can be used in the expression of GOIs in organisms, such as in *A. niger*.

The arabinofuranosidase of the present invention is different from the arabinofuranosidases previously known. In this regard, the previous described arabinofuranosidases - such as those of EP-A-0506190 - are characterised by their ability to degrade arabinan, and are assayed using p-nitrophenyl-arabinoside.

The arabinofuranosidase of the present invention does not degrade arabinan, and only a minor activity is seen on p-nitrophenyl-arabinoside.

25 In contrast, the arabinofuranosidase of the present invention is useful for degrading arabinoxylan. Therefore, the arabinofuranosidase of the present invention is quite different from the previous isolated arabinofuranosidases.

More in particular, the enzyme of the present invention is capable of specifically cleaving 30 arabinose from the xylose backbone of arabinoxylan.

The enzyme of the present invention is useful as it can improve processes for preparing foodstuffs and feeds as well as the foodstuffs and feeds themselves. For example, the enzyme of the present invention may be added to animal feeds which are rich in arabinoxylans. When added to feeds (including silage) for monogastric animals (e.g. 5 poultry or swine) which contain cereals such as barley, wheat, maize, rye or oats or cereal by-products such as wheat bran or maize bran, the enzyme significantly improves the break-down of plant cell walls which leads to better utilization of the plant nutrients by the animal. As a consequence, growth rate and/or feed conversion are improved. Moreover, arabinoxylan-degrading enzymes may be used to reduce the viscosity of feeds 10 containing arabinans. The arabinoxylan-degrading enzyme may be added beforehand to the feed or silage if pre-soaking or wet diets are preferred.

15 Of particular benefit is the use of the enzyme according to the present invention in combination with a xylanase, especially an endoxylanase.

A possible further application for the enzyme according to the present invention is in the pulp and paper industry. The application of xylanases is often reported to be beneficial in the removal of lignins and terpenoids from the cellulose and hemicellulose residues of a hemicellulose backbone, an essential step in the processing of wood, wood pulp or 20 wood derivative product for the production of paper. The addition of arabinoxylan-degrading enzymes, produced according to the present invention, to the xylanase treatment step should assist in the degradation of an arabinan-containing hemicellulose backbone and thus facilitate an improved, more efficient removal of both lignins and terpenoids. The application of arabinoxylan-degrading enzymes should be particularly 25 advantageous in the processing of soft woods in which the hemicellulose backbone contains glucuronic acid.

30 The enzyme according to the present invention is also useful as it acts in a synergistic manner with endoxylanase (see results presented above).

Other modifications of the present invention will be apparent to those skilled in the art without departing from the scope of the invention.

SEQUENCE LISTINGS

SEQ ID NO: 1

ENZYME SEQUENCE

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 296 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

Lys Cys Ser Leu Pro Ser
1 5

Ser Tyr Ser Trp Ser Ser Thr Asp Ala Leu Ala Thr Pro Lys Ser Gly
10 15 20

Trp Thr Ala Leu Lys Asp Phe Thr Asp Val Val Ser Asp Gly Lys His
25 30 35

Ile Val Tyr Ala Ser Thr Thr Asp Glu Ala Gly Asn Tyr Gly Ser Met
40 45 50

Thr Phe Gly Ala Phe Ser Glu Trp Ser Asn Met Ala Ser Ala Ser Lys
55 60 65 70

Thr Ala Thr Pro Tyr Asn Ala Val Ala Pro Thr Leu Phe Tyr Phe Lys
75 80 85

Pro Lys Ser Ile Trp Val Leu Ala Tyr Gln Trp Gly Ser Ser Thr Phe
90 95 100

Thr Tyr Arg Thr Ser Gln Asp Pro Thr Asn Val Asn Gly Trp Ser Ser
105 110 115

Glu Lys Ala Leu Phe Thr Gly Lys Leu Ser Asp Ser Ser Thr Gly Ala
120 125 130

Ile Asp Gln Thr Val Ile Gly Asp Asp Thr Asn Met Tyr Leu Phe Phe
135 140 145 150

Ala Gly Asp Asn Gly Lys Ile Tyr Arg Ser Ser Met Ser Ile Asp Glu
155 160 165

Phe Pro Gly Ser Phe Gly Ser Gln Tyr Glu Glu Ile Leu Ser Gly Ala
170 175 180
Thr Asn Asp Leu Phe Glu Ala Val Gln Val Tyr Thr Val Asp Gly Gly
185 190 195
Glu Gly Asn Ser Lys Tyr Leu Met Ile Val Glu Ala Ile Gly Ser Thr
200 205 210
Gly His Arg Tyr Phe Arg Ser Phe Thr Ala Ser Ser Leu Gly Gly Glu
215 220 225 230
Trp Thr Ala Gln Ala Ala Ser Glu Asp Lys Pro Phe Ala Ala Lys Pro
235 240 245
Thr Val Ala Pro Pro Gly Pro Lys Thr Leu Ala Met Val Thr Trp Phe
250 255 260
Ala Thr Thr Leu Ile Lys Pro *
265 270

SEQ ID NO: 2

NUCLEOTIDE CODING SEQUENCE

AAA TGC TCT CTT CCA TCG TCC TAT AGT TGG AGT TCA ACC GAT GCT CTC
GCA ACT CCT AAG TCA GGA TGG ACC GCA CTG AAG GAC TTT ACT GAT GTT
GTC TCT GAC GGC AAA CAT ATC GTC TAT GCG TCC ACT ACT GAT GAA GCG
GGA AAC TAT GGC TCG ATG ACC TTT GGC GCT TTC TCA GAG TGG TCG AAC
ATG GCA TCT GCT AGC AAG ACA GCC ACC CCC TAC AAT GCC GTG GCT CCT
ACC CTG TTC TAC TTC AAG CCG AAA AGC ATC TGG GTT CTG GCC TAC CAA
TGG GGC TCC AGC ACA TTC ACC TAC CGC ACC TCC CAA GAT CCC ACC AAT
GTC AAC GGC TGG TCG TCG GAG AAG GCG CTT TTC ACC GGA AAA CTC AGC
GAC TCA AGC ACC GGT GCC ATT GAC CAG ACG GTG ATT GGC GAC GAT ACG
AAT ATG TAT CTC TTC TTT GCT GGC GAC AAC GGC AAG ATC TAC CGA TCC
AGC ATG TCC ATC GAT GAA TTT CCC GGA AGC TTC GGC AGC CAG TAC GAG
GAA ATT CTG AGT GGT GCC ACC AAC GAC CTA TTC GAG GCG GTC CAA GTG
TAC ACG GTT GAC GGC GGC GAG AAC AGC AAG TAC CTC ATG ATC GTT
GAG GCG ATC GGG TCC ACT GGA CAT CGT TAT TTC CGC TCC TTC ACG GCC
AGC AGT CTC GGT GGA GAG TGG ACA GCC CAG GCG GCA AGT GAG GAT AAA
CCC TTC GCA GCA AAG CCA ACA GTG GCG CCA CCT GGA CCG AAG ACA TTA
GCC ATG GTG ACT TGG TTC GCA ACA ACC CTG ATC AAA CCA TGA

SEQ ID NO: 3
PROMOTER SEQUENCE

CTGCAGAAGA TGGCAGTCGC CACAGCCGAT CACCCGATCC ATACTGGATG TTGTAACCTTG 60
GAGACAGCCT GCAGATGCTC TGATGAAGGT CTGCAAATAG TTCTGGACC TCGATAAGTGA 120
AGTATAACCGA TTCGTCAATG TTGTATATCC AGCCACTTTG AAAGTACCAA CTTTAGTTC 180
GATTGATCAG AATACTTTG GTGTGTAACA TTGACAAGCC AAATTATCAA TCTCTTCTAC 240
CGGTAAGGTG TCAACTACCC GGCGAAAGT ACCGGAAGGT CGTGGTGTAA TAAGGTGAAA 300
CAACTATCAG GGCAGCAATG TGTCAAAGTA GAACCAGTTT GCTTAGCGCC ATTAGGATCC 360
.ACGCCTAGAC CCTTGATGCC CGGGAGTTAT CCGTCCTGTC ACAGCAATTA TTTCCCCGAG 420
TCTACTGCCG AAGAACAGCC ATTGTGGCGT ACTCACGGAA TTACCCACTG TGTAGGGTAG 480
TCTTGAACGC CGTTCTAGAC ACGGCAACGC TCCGGTGGAC GATCGTTTCT GGCTAATGTA 540
CTCCGTAGTT TAGGCAGCAT GCTGATCATC TTCCCCCTAG GGAAAGGCCCT CTGAATAGTG 600
CGCCAAAATG AGCTTGAGCA AAGGAATGTT CTTTCTAAGC CAAAGTGAGG GAAATAACCA 660
AGCAGCCCAC TTTTATCCGA AACGTTTCTG GTGTCACTCCA ATATGGATAA ATCCGATTG 720
TTCTTCTGCA CATATCTCTA TTGTCTAAG TGCAACTACA TATATTTGAA CATGGTTTGG 780
TCCTCTTCC AAGTTATTG TTCTCCGTGA CCAGCGATTT CAGCCATTGA TTCTTTGTT 840
TCTTTCCCCG CGGATAAACT CATACTGAAG

INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: N-terminal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

Lys Cys Ser Leu Pro Ser Ser Tyr Ser Trp Ser Ser Thr Asp Ala Leu
1 5 10 15
Ala Thr Pro Lys
20

INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 41 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

Tyr Leu Met Ile Val Glu Ala Ile Gly Ser Thr Gly His Arg Tyr Phe
1 5 10 15
Arg Ser Phe Thr Ala Ser Ser Leu Gly Gly Glu Met Thr Ala Gln Ala
20 25 30
Ala Ser Glu Asp Lys Pro Phe Xaa Gly
35 40

INFORMATION FOR SEQ ID NO: 6:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 25 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (v) FRAGMENT TYPE: internal
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
Ser Ile Trp Val Leu Ala Tyr Gln Trp Gly Ser Ser Thr Phe Thr Tyr
1 5 10 15
Arg Thr Ser Gln Asp Pro Thr Asn Val
20 25

INFORMATION FOR SEQ ID NO: 7:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (v) FRAGMENT TYPE: internal
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
Asp Ile Val Tyr Ala Ser Thr Thr Asp Glu Ala Gly Asn Tyr Gly Ser
1 5 10 15
Met Thr Phe Gly Ala Phe Ser Glu Xaa Ser Asn Met Ala Ser
20 25 30

INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 41 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

Ile Tyr Arg Ser Ser Met Ser Ile Asp Glu Phe Pro Gly Ser Phe Gly
1 5 10 15
Ser Gln Tyr Glu Glu Ile Leu Ser Gly Ala Thr Asn Asp Leu Phe Glu
 20 25 30
Ala Val Gln Val Tyr Thr Val Asp Gly
 35 40

INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "OLIGONUCLEOTIDE"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

GCYTGNGCNG TCATYTC

17

INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "OLIGONUCLEOTIDE"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

ATG ATH GTN GAR GCN ATH GG

20

INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 89 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "PCR fragment"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

ATGATTGTGG AGGCGATCGG GTCCACTGGA CATCGTTATT TCCGCTCCTT CACGGCCAGC	60
AGTCTCGGTG GAGAGATGAC CGCACAGGC	89

INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2555 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Aspergillus niger
- (B) STRAIN: 3M43

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 870..1757

(ix) FEATURE:

- (A) NAME/KEY: sig_peptide
- (B) LOCATION: 870..947

(ix) FEATURE:

- (A) NAME/KEY: mat_peptide
- (B) LOCATION: 948..1754

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

CTGCAGAAGA TGGCAGTCGC CACAGCCGAT CACCCGATCC ATACTGGATG TTGTAACCTG	60
GAGACAGCCT GCAGATGCTC TGATGAAGGT CTGCAAATAG TTCCCTGGACC TCGATAGTGA	120
AGTATAACCGA TTCGTCAATG TTGTATATCC AGCCACTTTG AAAGTACCAA CTTTAGTTC	180
GATTGATCAG AATACTTTG GTGTGTAACA TTGACAAGCC AAATTATCAA TCTCTCTAC	240

CGGTAAGGTG TCAACTACCC GGCGAAAGT ACCGGAAGGT CGTGGTGT TT TAAGGTGAAA	300
CAACTATCAG GGCGGCAATG TGTCAAAGTA GAACCAGTTT GCTTAGCGCC ATTAGGATCC	360
ACGCCTAGAC CCTTGATGCC CGGGAGTTAT CCGTCCTGTC ACAGCAATT A TTTCCCCGAG	420
TCTACTGCCG AAGAACAGCC ATTGTGGCGT ACTCACGGAA TTACCCACTG TGTAGGGTAG	480
TCTTGAACGC CGTTCTAGAC ACGGCAACGC TCCGGTGGAC GATCGTTCT GGCTAATGTA	540
CTCCGTAGTT TAGGCAGCAT GCTGATCATC TTCCCCCTAG GGAAAGGCCG CTGAATAGTG	600
CGCCAAAATG AGCTTGAGCA AAGGAATGTT CTTTCTAAGC CAAAGTGAGG GAAATAACCA	660
AGCAGCCCAC TTTTATCCGA AACGTTCTG GTGTATCCA ATATGGATAA ATCCCGATTG	720
TTCTTCTGCA CATATCTCTA TTGTCATAAG TGCAACTACA TATATTTGAA CATGGTTTGG	780
TCCTCTTCC AAGTTATTG TTCTCCGTGA CCAGCGATTT CAGCCATTGA TTCTTTTGT	840
TCTTCCCCG CGGATAAACT CATACTGAAG ATG AAG TTC TTC AAT GCC AAA GGC	893
Met Lys Phe Phe Asn Ala Lys Gly	
-26 -25 -20	
AGC TTG CTG TCA TCA GGA ATC TAC CTC ATT GCA TTA ACC CCC TTT GTT	941
Ser Leu Leu Ser Ser Gly Ile Tyr Leu Ile Ala Leu Thr Pro Phe Val	
-15 -10 -5	
AAC GCC AAA TGC TCT CTT CCA TCG TCC TAT AGT TGG AGT TCA ACC GAT	989
Asn Ala Lys Cys Ser Leu Pro Ser Ser Tyr Ser Trp Ser Ser Thr Asp	
1 5 10	
GCT CTC GCA ACT CCT AAG TCA GGA TGG ACC GCA CTG AAG GAC TTT ACT	1037
Ala Leu Ala Thr Pro Lys Ser Gly Trp Thr Ala Leu Lys Asp Phe Thr	
15 20 25 30	
GAT GTT GTC TCT GAC GGC AAA CAT ATC GTC TAT GCG TCC ACT ACT GAT	1085
Asp Val Val Ser Asp Gly Lys His Ile Val Tyr Ala Ser Thr Thr Asp	
35 40 45	
GAA GCG GGA AAC TAT GGC TCG ATG ACC TTT GGC GCT TTC TCA GAG TGG	1133
Glu Ala Gly Asn Tyr Gly Ser Met Thr Phe Gly Ala Phe Ser Glu Trp	
50 55 60	
TCG AAC ATG GCA TCT GCT AGC AAG ACA GCC ACC CCC TAC AAT GCC GTG	1181
Ser Asn Met Ala Ser Ala Ser Lys Thr Ala Thr Pro Tyr Asn Ala Val	
65 70 75	
GCT CCT ACC CTG TTC TAC TTC AAG CCG AAA AGC ATC TGG GTT CTG GCC	1229
Ala Pro Thr Leu Phe Tyr Phe Lys Pro Lys Ser Ile Trp Val Leu Ala	
80 85 90	
TAC CAA TGG GGC TCC AGC ACA TTC ACC TAC CGC ACC TCC CAA GAT CCC	1277
Tyr Gln Trp Gly Ser Ser Thr Phe Thr Tyr Arg Thr Ser Gln Asp Pro	
95 100 105 110	

ACC AAT GTC AAC GGC TGG TCG TCG GAG GCG CTT TTC ACC GGA AAA		1325	
Thr Asn Val Asn Gly Trp Ser Ser Glu Lys Ala Leu Phe Thr Gly Lys			
115	120	125	
CTC AGC GAC TCA AGC ACC GGT GCC ATT GAC CAG ACG GTG ATT GGC GAC		1373	
Leu Ser Asp Ser Ser Thr Gly Ala Ile Asp Gln Thr Val Ile Gly Asp			
130	135	140	
GAT ACG AAT ATG TAT CTC TTC TTT GCT GGC GAC AAC GGC AAG ATC TAC		1421	
Asp Thr Asn Met Tyr Leu Phe Phe Ala Gly Asp Asn Gly Lys Ile Tyr			
145	150	155	
CGA TCC AGC ATG TCC ATC GAT GAA TTT CCC GGA AGC TTC GGC AGC CAG		1469	
Arg Ser Ser Met Ser Ile Asp Glu Phe Pro Gly Ser Phe Gly Ser Gln			
160	165	170	
TAC GAG GAA ATT CTG AGT GGT GCC ACC AAC GAC CTA TTC GAG GCG GTC		1517	
Tyr Glu Glu Ile Leu Ser Gly Ala Thr Asn Asp Leu Phe Glu Ala Val			
175	180	185	190
CAA GTG TAC ACG GTT GAC GGC GGC GAG GGC AAC AGC AAG TAC CTC ATG		1565	
Gln Val Tyr Thr Val Asp Gly Glu Gly Asn Ser Lys Tyr Leu Met			
195	200	205	
ATC GTT GAG GCG ATC GGG TCC ACT GGA CAT CGT TAT TTC CGC TCC TTC		1613	
Ile Val Glu Ala Ile Gly Ser Thr Gly His Arg Tyr Phe Arg Ser Phe			
210	215	220	
ACG GCC AGC AGT CTC GGT GGA GAG TGG ACA ACA GCG CAG GCG GCA AGT GAG		1661	
Thr Ala Ser Ser Leu Gly Glu Trp Thr Ala Gln Ala Ala Ser Glu			
225	230	235	
GAT AAA CCC TTC GCA GCA AAG CCA ACA GTG GCG CCA CCT GGA CGG AAG		1709	
Asp Lys Pro Phe Ala Ala Lys Pro Thr Val Ala Pro Pro Gly Pro Lys			
240	245	250	
ACA TTA GCC ATG GTG ACT TGG TTC GCA ACA ACC CTG ATC AAA CCA TGA		1757	
Thr Leu Ala Met Val Thr Trp Phe Ala Thr Thr Leu Ile Lys Pro *			
255	260	265	270
CTGTCGATCC TTGCAACCTC CAGTTGCTCT ATCAGGGCCA TGACCCCCAA CAGCAGTGGC		1817	
GACTACAACC TCTTGCCATG GAAGCCGGGC GTCCTTACCT TGAAGCAGTG ACGAGCTTAT		1877	
CTTTAGTTGC AGATCGTGTT TCTCCTTCT TCTTCAAGTA GTTTTAGTGG TGGAAAGACAG		1937	
CAGAAGGTGG TCATCATCTT AGGCTCAGTT GGGGTGGGCT CCTGCCACGT TTTGTCCATA		1997	
GGCTAGTAAT TTGCACGGAA TTCAGTTCAT TGGCAAGGAG TGCAGTACGA ATACCTGTTT		2057	
TCACAATAGC AATTAGGCCA AGTAGTTATA CTACGTACTG GAATTGAGTA CTCGTAGTAG		2117	
CAAGATTGTT TGCCTCAGAG GGAATGGCCG ACACGTGAGC AAGTCACCTT CATCAGCTAG		2177	

TCGCGTCCA CATAGACAAT GGTCCAGCTC CAGAGTGGAA TTTGGGCTAC TTTGAACGAT	2237
GGCCGATTGA ATCGCGCGTC TCCTCAATTG TATTTAACCA CAATAGGCCA GGTATTGGCA	2297
TTCACTCTCC GCCTTGCGG GTGCCGGCAC GAGATGTCTC CTGAAGAAC TAGGCAACGA	2357
GCAGACTGTG GATATGGGAG ATGGTTGACG ATGTGCTTCT TGGTAAATT GAAGCCTCCA	2417
GGGCCTCTAG AAAGGCGGGA ATTAAATCT CAAGTGCCT AACGTGTCCG ACCACGGTGT	2477
TGATCATCAT TCATTGAATC GGATAACAGT CTTGGTTCGG AACTGAACA GGCGGCTCTT	2537
GAATGACACT CTGGATCC	2555

(2) INFORMATION FOR SEQ ID NO: 13:

TERMINATOR SEQUENCE

CTGTCGATCC TTGCAACCTC CAGTTGCTCT ATCAGGGCCA TGACCCCCAA CAGCAGTGGC	60
GACTACAACC TCTTGCCATG GAAGCCGGGC GTCCCTTACCT TGAAGCAGTG ACGAGCTTAT	120
CTTAGTTGC AGATCGTGTT TCTCCTTCT TCTTCAAGTA GTTTTAGTGG TGGAAAGACAG	180
CAGAAGGTGG TCATCATCTT AGGCTCAGTT GGGGTGGGCT CCTGCCACGT TTTGCCATA	240
GGCTAGTAAT TTGCACGGAA TTCAGTTCAT TGGCAAGGAG TGCGGTACGA ATACCTGTTT	300
TCACAATAGC AATTAGGCCA AGTAGTTATA CTACGTACTG GAATTGAGTA CTCGTAGTAG	360
CAAGATTGTT TGCCTCAGAG GGAATGGCCG ACACGTGAGC AAGTCACCTT CATCAGCTAG	420
TCGCGTCCA CATAGACAAT GGTCCAGCTC CAGAGTGGAA TTTGGGCTAC TTTGAACGAT	480
GGCCGATTGA ATCGCGCGTC TCCTCAATTG TATTTAACCA CAATAGGCCA GGTATTGGCA	540
TTCACTCTCC GCCTTGCGG GTGCCGGCAC GAGATGTCTC CTGAAGAAC TAGGCAACGA	600
GCAGACTGTG GATATGGGAG ATGGTTGACG ATGTGCTTCT TGGTAAATT GAAGCCTCCA	660
GGGCCTCTAG AAAGGCGGGA ATTAAATCT CAAGTGCCT AACGTGTCCG ACCACGGTGT	720
TGATCATCAT TCATTGAATC GGATAACAGT CTTGGTTCGG AACTGAACA GGCGGCTCTT	780
GAATGACACT CTGGATCC	798

(2) INFORMATION FOR SEQ ID NO: 14

Signal SEQUENCE

ATG AAG TTC TTC AAT GCC AAA GGC AGC TTG CTG TCA TCA GGA ATC TAC	48
CTC ATT GCA TTA ACC CCC TTT GTT AAC GCC	78

SEQ ID NO: 15

SIGNAL SEQUENCE

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 26 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

Met Lys Phe Phe Asn Ala Lys Gly Ser Leu	10
Leu Ser Ser Gly Ile Tyr Leu Ile Ala Leu	20
Thr Pro Phe Val Asn Ala	26

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description
on page 27, line 5 25 and 26

B. IDENTIFICATION OF DEPOSIT

Further deposits are identified on an additional sheet:

Name of depository institution

The National Collections of Industrial and Marine Bacteria Limited (NCIMB)

Address of depository institution (including postal code and country)

23 St. Machar Drive
Aberdeen
Scotland
AB2 1RY
United Kingdom

Date of deposit

16 JANUARY 1995

Accession Number

NCIMB 40703

C. ADDITIONAL INDICATIONS (Leave blank if not applicable) This information is continued on an additional sheet:

In respect of those designations in which a European patent is sought, and any other designated state having equivalent legislation, a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which the application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample. (Rule 16(4) EPC).

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)

E. SEPARATE FURNISHING OF INDICATIONS (Leave blank if not applicable)

The indications listed below will be submitted to the International Bureau later, specifying the nature of the indications e.g., "Classification Number of Deposit".

For receiving Office use only

This sheet was received with the international application

Authorized officer:

J. van Aube

For International Bureau use only

This sheet was received by the International Bureau on

Authorized officer:

CLAIMS

1. An enzyme that is obtainable from *Aspergillus*, wherein the enzyme has the following characteristics:

5

- a. a MW of 33,270 D \pm 50 D
- b. a pI value of about 3.7
- c. arabinoxylan degrading activity
- d. a pH optima of from about 2.5 to about 7.0 (more especially from about 10 3.3 to about 4.6, more especially about 4)
- e. a temperature optima of from about 40°C to about 60°C (more especially from about 45°C to about 55°C, more especially about 50°C);

10

wherein the enzyme is capable of cleaving arabinose from the xylose backbone of an 15 arabinoxylan.

2. An enzyme having the sequence shown as SEQ. I.D. No. 1 or a variant, homologue or fragment thereof.

20

3. An enzyme coded by the nucleotide sequence shown as SEQ. I.D. No. 2 or a variant, homologue or fragment thereof or a sequence complementary thereto.

4. A nucleotide sequence coding for the enzyme according to claim 1.

25

5. A nucleotide sequence coding for the enzyme according to claim 2.

6. A nucleotide sequence having the sequence shown as SEQ. I.D. No. 2 or a variant, homologue or fragment thereof or a sequence complementary thereto.

30

7. A nucleotide sequence according to any one of claims 4 to 6 operatively linked to a promoter.

8. A nucleotide sequence according to claim 7 wherein the promoter is the promoter having the sequence shown as SEQ. I.D. No. 3 or a variant, homologue or fragment thereof or a sequence complementary thereto.
- 5 9. A promoter having the sequence shown as SEQ. I.D. No. 3 or a variant, homologue or fragment thereof or a sequence complementary thereto.
10. A promoter according to claim 9 operatively linked to a GOI.
- 10 11. A promoter according to claim 10 wherein the promoter is operatively linked to a GOI, wherein the GOI comprises a nucleotide sequence according to any one of claims 4-6.
- 15 12. A terminator having the nucleotide sequence shown as SEQ. I.D. No. 13 or a variant, homologue or fragment thereof or a sequence complementary thereto.
13. A signal sequence having the nucleotide sequence shown as SEQ. I.D. No. 14 or a variant, homologue or fragment thereof or a sequence complementary thereto.
- 20 14. A construct comprising or expressing the invention according to any one of claims 1 to 13.
15. A vector comprising or expressing the invention of any one of claims 1 to 14.
- 25 16. A plasmid comprising or expressing the invention of any one of claims 1 to 15.
17. A transgenic organism comprising or expressing the invention according to any one of claims 1 to 16.
- 30 18. A transgenic organism according to claim 17 wherein the organism is a fungus.

19. A transgenic organism according to claim 18 wherein the organism is a filamentous fungus, preferably of the genus *Aspergillus*.

20. A transgenic organism according to claim 17 wherein the organism is a plant.

5

21. A process of preparing an enzyme according to any one of claims 1 to 3 comprising expressing a nucleotide sequence according to any one of claims 4-8.

22. A process according to claim 21 wherein the enzyme has the sequence shown as SEQ. I.D. No. 1 or a variant, homologue or fragment thereof, and the nucleotide sequence has the sequence shown as SEQ. I.D. No. 2 or a variant, homologue or fragment thereof or a sequence complementary thereto.

10 23. A process according to claim 21 or claim 22 wherein the expression is controlled
15 (partially or completely) by use of a promoter according to claim 9.

24. A process for expressing a GOI by use of a promoter, wherein the promoter is the promoter according to claim 9.

20 25. Use of an enzyme according to any one of claims 1 to 3 or prepared by a process according to any one of claims 21 to 24 to degrade an arabinoxylan.

26. Use according to claim 24 wherein the enzyme is used in combination with a xylanase, preferably an endoxylanase.

25

27. A combination of enzymes to degrade an arabinoxylan, the combination comprising an enzyme according to any one of claims 1 to 3 or prepared by a process according to any one of claims 21 to 24 claims; and a xylanase.

30 28. Plasmid NCIMB 40703, or a nucleotide sequence obtainable therefrom for expressing an enzyme capable of degrading arabinoxylan or for controlling the expression thereof or for controlling the expression of another GOI.

29. A signal sequence having the sequence shown as SEQ. I.D. No. 15 or a variant, homologue or fragment thereof.

30. The use of the enzyme according to any one of claims 1 to 3 or prepared by a process according to any one of claims 21 to 24 claims, in the manufacture of a medicament or foodstuff to reduce or prevent indigestion and/or increase nutrient absorption.

31. An arabinofuranosidase enzyme having arabinoxylan degrading activity, which is immunologically reactive with an antibody raised against a purified arabinofuranosidase enzyme having the sequence shown as SEQ. I.D. No. 1.

FIGURE 1

AMY 637 PROMOTER

SEQUENCE TYPE: Nucleotide

MOLECULE TYPE: DNA

ORIGINAL SOURCE: *Solanum Tuberosum*

SEQUENCE LENGTH: 2094

SEQUENCE:

10	20	30	40
ATTAAGGGGA	GCATAAGTGC	AGCTCAGAAA	TTCACACCTG
50	60	70	80
ATATTTCCC	AAAGCCCTCA	AAAATGTGAA	CAAATCTGCT
90	100	110	120
AAAATGTCA	TCAGAAGGAC	TGTTCTTTA	GGTTTTCTTC
130	140	150	160
TCTCGAGTCA	CGAAATCAGA	TAATATGATA	AGAAATTATG
170	180	190	200
GAGGATTAT	AATGTATCTG	TCTGTTCTTA	GGTATAATTA
210	220	230	240
TGTGTTCCCTT	TATGATGTAG	TAATGGAATT	CTGGGCTTAT
250	260	270	280
ATAAAGGAA	CTGAATATAA	ATGTTCGCAT	TTAACTGCG
290	300	310	320
GAGACTTCA	GTTAGAGCCT	TATAATTATG	TCTTATCATT
330	340	350	360
TTAACTGAG	ATCATATTAC	AGATGATGAA	AGCTGACATT
370	380	390	400
GCATTAGTTA	TTCTGTTTA	TACAAGTCAT	GTAAC TGCTG
410	420	430	440
CTTGTGAGTT	GTGACTGTAA	GATAAATTGA	TTCAGCCTTC
450	460	470	480
TGTGGCATT	GC GGAGATCT	GATTATACTC	TCATCGTCTT
490	500	510	520
ATCTAAGTTG	CTCATGCAAC	TTTGTCCCTTG	ATAGTTGGCT
530	540	550	560
AATACTACAA	CTGGAATTAA	GTGTAGTTAT	TCGAAATCTC
570	580	590	600
TGTTGGAAGT	TGCTAAGTGC	TTAAGTGCTG	GTTATTGTAA
610	620	630	640
ACCCCATCCG	AGTTATTATA	CAGCATCTGG	CTGATGAAAT
650	660	670	680
GCTGCTCATT	TGCAATGGTG	ACATAACCAA	ATGTTAGTAA
690	700	710	720
AACATACTAG	CTGGTTGAAT	GTTAGATGAT	TGTTCAACGT
730	740	750	760
TACATCTCAC	AGAAACCTTA	TTATGGATTG	ACATGTTAGT

FIGURE 1 CONTINUED

770	780	790	800
TGATCCGAAA	GATCCTTCTT	TTAAATGCCA	AAGCTTGTTA
810	820	830	840
CAGATTTGAG	GAGTTCTTTT	ACTTTCTTTT	GTATATCTA
850	860	870	880
TTTCCCCATTTC	ATTTTGACGT	TCAGCCTCAC	AGATGTTGTC
890	900	910	920
ATACTTAGAA	ATGTGGTAT	ATATATAGAG	AGAGAGAGAT
930	940	950	960
AGAGTGAAAT	GATTATATAG	TCSAAGATTA	CGAAACTTGA
970	980	990	1000
CATTGAGACA	TCTGTGATTG	TTTGAATTT	ATGTATATAT
1010	1020	1030	1040
CTGTAGCATT	AGAAAACATA	AGAGTTGTTA	GCTTCACTTG
1050	1060	1070	1080
TCTTATTGTT	GTGCTCAAAG	CAACTTCATC	ATACAGTATG
1090	1100	1110	1120
GTTTTTATAT	GCTCTTCCAT	TATCACCGAA	CCTTATGATT
1130	1140	1150	1160
ATGTGTACGA	GCTTATAATA	TTACTGATGG	TGATTCACTA
1170	1180	1190	1200
TTATGATTAT	GTCCTCCATT	AATTATTCTG	TTTCATACAA
1210	1220	1230	1240
GTCGTGTAAT	TTGCTGTTG	TGATTGTACG	ATAAATTGAT
1250	1260	1270	1280
TCAACCTTCT	GCGGTGTTGG	TTGAAGTTCA	AGTAAATTAG
1290	1300	1310	1320
CTTTTATTAT	CATAGTAGCA	TTTGATTATT	GATGCTCTGT
1330	1340	1350	1360
AGCTAATGAT	AAGCCATTGA	AGGGAAAGCAG	AAATGGTAAA
1370	1380	1390	1400
GCTTTCTAAA	ATGAATCTAC	GAATGGATGA	TAAAGTTAAT
1410	1420	1430	1440
GAATATTGTT	GATACTTCTG	CAATCAGATT	ATGAGTTACT
1450	1460	1470	1480
GAGTCTACTG	TTTTTTAACG	CTGTTTCAGA	TGATCGATCA
1490	1500	1510	1520
TCAACAAACAA	CATATTCACT	GTAGTAGACA	TGATCGATCA
1530	1540	1550	1560
CTTTCTAATT	TTCGATTATG	CACCCCTTTT	TCTCCAATT
1570	1580	1590	1600
GGTCGTCTTC	TTTTTTTCAT	GATGTCACTG	AATTATTCTC
1610	1620	1630	1640
TGGTCGTCCC	CACCATTCAAG	GAAGTCACTT	CGAGCATAAT
1650	1660	1670	1680
GTGAAAACAT	CCACATTTT	CAAATCCAGC	AGAATTTCAT

FIGURE 1 CONTINUED

3/38

1690 1700 1710 1720
TCAAACGGGG TTCAACATTT ACTACATGTA TACACTCTGA
1730 1740 1750 1760
AGTCTGAATC CACTAATTCT AGATGGTGCA TCTGTGCCCG
1770 1780 1790 1800
CACACTTG TGAAAGCTTATT CTCAATTTCCTT TATTTTCCAA
1810 1820 1830 1840
CAACTTGAAAT TCAGACCACA CAACTCCCGT GTCTTGTACG
1850 1860 1870 1880
GTCAGCATCT GAGTGGAGAA CTCAATTAAAG TGACTTTAAC
1890 1900 1910 1920
GTCGAGTTCT ATAGTAAACA ACCCCTATAT CTTTTTCAA
1930 1940 1950 1960
GCATGTTAAG ATTGCGAACCA CACTGAAATT TCCAGGTCGT
1970 1980 1990 2000
TAATCTTGTAA CCCAGTGTGT GTACTTTAA AAAAAAAAAGT
2010 2020 2030 2040
CAGTTTTTTA GTCTCTAAAA CACATTTAAA TAGAGTTTAT
2050 2060 2070 2080
TTGCCATCTT TTGTTCCCTCA TACTAGACTT CGGAGTCAAC
2090
ACAAACACAAAC AACAA

FIGURE 2

AMY 351 PROMOTER

SEQUENCE TYPE: Nucleotide

MOLECULE TYPE: DNA (genomic)

ORIGINAL SOURCE: *Solanum tuberosum*

SEQUENCE LENGTH: 1734 bp

STRANDEDNESS: Double

TOPOLOGY: Linear

SEQUENCE:

TCTTTAAGT	10	GTTGCTTGA	20	TTCCTTCT	30	TCAATCTTCT	40
ATATTAAAT	50	CGTTTAGCT	60	TCAAACCTCT	70	TCAATTTAT	80
TTCAATTAA	90	TTCTACAAAAA	100	AAAATCTCTA	110	TTAGCACCA	120
TTCATAAAAT	130	TCATGCTCAA	140	AATGGGCAA	150	CATAAATAAT	160
AAATGTGAAG	170	TAAATAATGG	180	ATTAAAATAT	190	ATATTTTGG	200
GCCTCACATC	210	AACCTTCATA	220	ATTCTTGAAT	230	GAATGAATGA	240
TAGACTTCAT	250	AATTTTTAA	260	CCTATACATA	270	TAAGAAAATT	280
GAGAGTAAC	290	CAAATAACAA	300	GTTGTAGTAT	310	CACATCTTA	320
CTATTTGATA	330	ACATTATGAA	340	GGTGATTATA	350	CATTACGTAA	360
CATTTCTTT	370	AAAAATATGT	380	AAGCAAATT	390	ACTTTTAAC	400
TTATCATTGA	410	TCTTCATGGT	420	TTTGTATAA	430	ATCTCAAAGT	440
TATCATATTT	450	TATATAGCTA	460	TTTGAAAGTA	470	ATTTTATTT	480
TACTCATCAT	490	TGAGTGATGC	500	TTTATTATA	510	ATACTAGTAA	520
GTTTTATTTA	530	TTATTTCTT	540	TTAGGGGTGA	550	ATTGTATAAT	560
ATAATAAAAAA	570	ATATTTTTT	580	590	590	600	600
TATTAAGATATT	610	AGAAATAATG	620	630	630	640	640
ATAATGAACG	650	AGATTATTTA	660	TGCTTGATA	670	680	680
ATGTTTGTAA	690	AAGTTTATT	700	TTCTATGAGT	710	TTCTTATAAT	720
TTTGTATTAT	730	TTATTTCAA	740	TTTGATGTA	750	TTTTTATAAT	760
TATATTATTA		TACTATATT		AAAAATTAA			

FIGURE 2 CONTINUED

770	780	790	800
AGATCCATAG	GGCTTACGCC	CCACGTCAAG	AGGCTTGC
810	820	830	840
CTTTCCCTAA	ATTAAGTAAA	ACTCTTCGCC	TGATGCC
850	860	870	880
CGCCTCGGCC	TTTAAACACA	CTGATTCC	TGCTCATATA
890	900	910	920
GCTTGAGGCG	AAAATATTTA	ATAAAAAACAC	TTCTTAATT
930	940	950	960
GTTTATATGT	TCAATTGAAC	ATGTCG	TAGAAAATT
970	980	990	1000
AAATTAAATT	CAATGACAAA	TTAATAATT	TGACACAAAA
1010	1020	1030	1040
TTTATGAAAA	AAATATCAAA	ATATAAAGAA	ATATTTTT
1050	1060	1070	1080
TGAAATGGAT	AAAAAAGAAA	AAAAAAACAA	ATAAATTGAA
1090	1100	1110	1120
CCGGGATAAG	TTGGTTGTTT	AATTGATTAT	TGATTATGAT
1130	1140	1150	1160
CTCAATTGGA	CATTTTGC	GATCTTCGA	CCTCAATTG
1170	1180	1190	1200
TATGAACTGA	CACTACGCCA	ATGGACAGTC	GCCGTCGTCA
1210	1220	1230	1240
CCGCCACCGC	ACTATTCTG	ACGCGTCGTC	TATCTCCTCC
1250	1260	1270	1280
ACCCCCACAGC	CGTCAATTCC	AAGCTTCAA	TGAACCGTTG
1290	1300	1310	1320
CCATGTGTCA	CTGCCTATT	ACCGCGAAC	ATGAATATCA
1330	1340	1350	1360
CTGACGAACG	ATTCGGAGC	GGAACGAATC	CAGAAAATGG
1370	1380	1390	1400
ATTACTTTCT	ATAAATTCT	CGAATCTCAA	CTCCATTTCG
1410	1420	1430	1440
TAAAAAATAAA	ATTAAAAATA	TTGTTTCTT	TTGTATTCT
1450	1460	1470	1480
TTTTGTATT	CTGGTTTATG	TGGTGATCGA	ATTTTCAATT
1490	1500	1510	1520
TTTTTACTGG	TAGTGATTCC	TACTTTTCTT	CAATTGCATT
1530	1540	1550	1560
TCTCCTTTT	CCATTTCACG	GTTGAGAATT	CATGATTCT
1570	1580	1590	1600
TATCAGAGGA	ATCGATCCGA	TTTGACTAAT	TTCAC
1610	1620	1630	1640
GTCTGTATAA	ATACCAGAGT	ATCTAGGTTG	AGGAACGTAA
1650	1660	1670	1680
TTTCAAGCTG	CGATCGGCTT	TTTCCCCTGA	ACGAGCAAAC

6/38

FIGURE 2 CONTINUED

1690 1700 1710 1720
ACAGGTTGTG GGTTCAGTT AGCAAGGGAC GTATAATCTC
1730
AACTACAATC CATT

7/38

FIGURE 3

 α -AMYLASE CODING SEQUENCE

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2017 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

- (A) LENGTH: 475 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

ATG AAG TCT CTC GCC GCA ATT GCT GCT CTG CTG TCG CCC ACA CTG GTC	48
Met Lys Ser Leu Ala Ala Ile Ala Ala Leu Leu Ser Pro Thr Leu Val	
-18 -15 -10 -5	
CGG GCA GCG ACT CCG GAT GAG TGG AAA GCT CAG TCG ATC TAT TTC ATG	96
Arg Ala Ala Thr Pro Asp Glu Trp Lys Ala Gin Ser Ile Tyr Phe Met	
1 5 10	
CTG ACG GAC CGG TTT GCG CGT ACC GAC AAT TCG ACC ACG GCT CCC TGT	144
Leu Thr Asp Arg Phe Ala Arg Thr Asp Asn Ser Thr Thr Ala Pro Cys	
15 20 25 30	
GAC ACC ACT GCC GGG GTATGCAACT AACCCCTGTGT TTCTCTTCCC GGGACGTACA	199
Asp Thr Thr Ala Gly	
35	
AGGGGTCTTC TCCATGCTAA CCGTGCACAT GCAG AAA TAT TGC GGG GGA ACA	251
Lys Tyr Cys Gly Gly Thr	
40	
TGG CGA GGT ATC ATC AAC AAC GTAAGTGGCT TCTGATTTTC GCTCAATAAT	302
Trp Arg Gly Ile Ile Asn Asn	
45	
CTTCGTCGCG TGACTTTATT TCCTAG CTG GAT TAC ATC CAG GAT ATG GGC TTC	355
Leu Asp Tyr Ile Gin Asp Met Gly Phe	
50 55	
ACA GCT ATC TGG ATA ACT CCA GTG ACA GCC CAG TGG GAC GAC GAT GTG	403
uThr Ala Ile Trp Ile Thr Pro Val Thr Ala Gin Trp Asp Asp Asp Val	
60 65 70	
GAT GCG GCA GAT GCA ACG TCG TAT CAC GGT TAT TGG CAG AAA GAC CT	450
Asp Ala Ala Asp Ala Thr Ser Tyr His Gly Tyr Trp Gin Lys Asp Leu	
75 80 85	
GTGCGCAACC CTGCTCCATG GATCGCTGGC TGCAAACTCG TGCTGATCGG TGATTTTTT	510
TTTTTTTTTT TTGAAACAG A TAC TCT CTG AAT TCG AAA TTC GGC ACT GCC	560
Tyr Ser Leu Asn Ser Lys Phe Gly Thr Ala	
90 95	

8/38

FIGURE 3 CONTINUED

GAT GAC TTG AAA GCC CTG GCT GAT GCC CTT CAC GCC CGT GGG ATG CTT Asp Asp Leu Lys Ala Leu Ala Asp Ala Leu His Ala Arg Gly Met Leu 100 105 110 115	608
CTC ATG GTC GAC GTC GTG GCT AAT CAC TTT GTACGGACCA TGTACATACCG Leu Met Val Asp Val Val Ala Asn His Phe 120 125	658
TGGGAAACGC GAAGAAGGAA AAAAAAAAAGGCGCACGC TAACATTTCG CGTTTAG 715	
GGC TAC GGC GGT TCT CAT AGC GAG GTG GAT TAC TCG ATC TTC AAT CCT Gly Tyr Gly Ser His Ser Glu Val Asp Tyr Ser Ile Phe Asn Pro 130 135 140	763
CTG AAC AGC CAG GAT TAC TTC CAC CCG TTC TGT CTC ATT GAG GAC TAC Leu Asn Ser Gln Asp Tyr Phe His Pro Phe Cys Leu Ile Glu Asp Tyr 145 150 155	811
GAC AAC CAG GAA GAA GTC GAA CAA TGC TGG CTG GCC GAT ACT CCG ACG Asp Asn Gln Glu Glu Val Glu Gln Cys Trp Leu Ala Asp Thr Pro Thr 160 165 170	859
ACA TTG CCC GAC GTG GAC ACC ACC AAT CCT CAG GTT CGG ACG TTT TTC Thr Leu Pro Asp Val Asp Thr Thr Asn Pro Gln Val Arg Thr Phe Phe 175 180 185	907
AAC GAC TGG ATC AAG AGC CTG GTG GCG AAC TAC TCC A GTATGATTGT Asn Asp Trp Ile Lys Ser Leu Val Ala Asn Tyr Ser 190 195 200	954
TCCCGCGGTA ACGCTTTAGG GCTTGCTCTA ACTGAAATCG ACAG TC GAT GGT CTG Ile Asp Gly Leu 205	1009
CGC GTC GAC ACC GTT AAG CAC GTG GAG AAA GAT TTC TGG CCC GAC TTC Arg Val Asp Thr Val Lys His Val Glu Lys Asp Phe Trp Pro Asp Phe 210 215 220	1057
AAC GAA GCT GCT GCG TGT ACC GTC GGC GAG GTG TTC AAC GGT GAC CCA Asn Glu Ala Ala Cys Thr Val Gly Glu Val Phe Asn Gly Asp Pro 225 230 235	1105
GCG TAC ACC TGC CCA TAC CAG GAA GTG CTG GAT GGC GTT CTG AAC TAT Ala Tyr Thr Cys Pro Tyr Gln Glu Val Leu Asp Gly Val Leu Asn Tyr 240 245 250	1153
CCG AT GTGAGTGATT CCGAAAGTTC CATCGATCAG GCTTTCTGAC GCATGAGAAC Pro Ile 255	1208

9/38

FIGURE 3 CONTINUED

AGC TAC TAT CCT GCG CTT GAT GCA TTC AAG TCT GTC GGC GGC AAT CTC Tyr Tyr Pro Ala Leu Asp Ala Phe Lys Ser Val Gly Gly Asn Leu 260 265 270	1256
GGC GGC TTG GCT CAG GCC ATC ACC ACC GTG CAG GAG AGC TGC AAG GAT Gly Gly Leu Ala Gln Ala Ile Thr Thr Val Gln Glu Ser Cys Lys Asp 275 280 285	1304
TCC AAT CTG CTC GGC AAT TTC CTT GAG AAT CAC GAC ATT GCT CGC T Ser Asn Leu Gly Asn Phe Leu Glu Asn His Asp Ile Ala Arg Phe 290 295 300	1352
GCT TC GTATGGACAC TCTTTTGAA GCCCTCATCG ATTGGGGATG CTGACACGGAA Ala Ser	1407
CAACAACAAAC AG G TAC ACG GAT GAC CTT GCT CTC GCC AAG AAT GGT CTC Tyr Thr Asp Asp Leu Ala Leu Ala Lys Asn Gly Leu 305 310 315	1456
GCT TTC ATC ATC CTC TCG GAT GGT ATT CCG ATC ATC TAC ACG GGC CAG Ala Phe Ile Ile Leu Ser Asp Gly Ile Pro Ile Ile Tyr Thr Gly Gln 320 325 330	1504
GAG CAG CAC TAC GCC GGT GAT CAC GAT CCC ACA AAT CGT GAG GCC GTC Glu Gln His Tyr Ala Gly Asp His Asp Pro Thr Asn Arg Glu Ala Val 335 340 345	1552
TGG CTG TCT GGC TAC AAT ACC GAC GCC GAG CTG TAC CAG TTC ATC AAG Trp Leu Ser Gly Tyr Asn Thr Asp Ala Glu Leu Tyr Gln Phe Ile Lys 350 355 360	1600
AAG GCC AAT GGC ATC CGC AAC TTG GCT ATC AGC CAG AAC CCG GAA TTC Lys Ala Asn Gly Ile Arg Asn Leu Ala Ile Ser Gln Asn Pro Glu Phe 365 370 375 380	1648
ACC TCC TCC AAG GTGAGTACAA TAACAAACTT TTTCGAAAAAT TTTTCACCGG Thr Ser Ser Lys	1700
AGAAAACCTA AGATTGGCT AACAAAACAA AAAAAAAAAA G ACC AAG GTC ATC Thr Lys Val Ile 385	1753
TAC CAA GAC GAT TCG ACC CTT GCC ATT AAC CGG GGC GGC GTC GTT ACT Tyr Gln Asp Asp Ser Thr Leu Ala Ile Asn Arg Gly Gly Val Val Thr 390 395 400	1801
GTC CTG AGC AAT GAA GGC GCC TCC GGG GAG ACC GGG ACT GTC TCC ATT Val Leu Ser Asn Glu Gly Ala Ser Gly Glu Thr Gly Thr Val Ser Ile 405 410 415 420	1849

10/38

FIGURE 3 CONTINUED

CCG GGA ACT GGC TTC GAG GCG GGC ACG GAA TTG ACT GAT GTC ATC TCC Pro Gly Thr Gly Phe Glu Ala Gly Thr Glu Leu Thr Asp Val Ile Ser	1897
425 430 435	
TGC AAG ACC GTG ACT GCG GGG GAC AGC GGG GCG GTC GAC GTG CCC TTG Cys Lys Thr Val Thr Ala Gly Asp Ser Gly Ala Val Asp Val Pro Leu	1945
440 445 450	
TCG GGC GGA CTG CCA AGC GTG CTC TAT CCC AGC TCC CAG CTG GCC AAG Ser Gly Leu Pro Ser Val Leu Tyr Pro Ser Ser Gln Leu Ala Lys	1993
455 460 465	
AGT GGT CTG TGT GCG TCG GCG TGA Ser Gly Leu Cys Ala Ser Ala	2017
470 475	

FIGURE 4

11/38

α -AMYLASE CODING SEQUENCE
 SEQUENCE TYPE: Nucleotide
 MOLECULE TYPE: DNA
 ORIGINAL SOURCE: *Solanum Tuberosum*
 SEQUENCE LENGTH: 1570
 SEQUENCE:

10	20	30	40
TGTGGTGATC	GAATTTC	CAA	TTT
50	60	70	80
GTTGAGGAAAC	GTAATTCAA	GCTGCAGATCG	GCTTTTTCCC
90	100	110	120
CTGAACGAGC	AAACACAGGT	TGTGGGTTCG	AGTTAGCAAG
130	140	150	160
GGACGTATAA	TCTCAACTAC	AATCCATTAT	GGCGCTTGAT
170	180	190	200
GAAAGTCAGC	AGTCTGATCC	ATTGGTTGTG	ATACGCAATG
210	220	230	240
GAAAGGAGAT	CATATTGCAG	GCATTGACT	GGGAATCTCA
250	260	270	280
TAAACATGAT	TGGTGGCTAA	ATTTAGATAC	GAAAGTTCT
290	300	310	320
GATATTGCAA	AGTCTGGTTT	CACAAGTGC	TGGCTGCCTC
330	340	350	360
CGGTGTGTCA	GTCATTGGCT	CCTGAAGGTT	ACCTTCCACA
370	380	390	400
GAACCTTTAT	TCTCTCAATT	CTAAATATGG	TTCTGAGGAT
410	420	430	440
CTCTTAAAAG	CTTTACTTAA	TAAGATGAAG	CAGTACAAAG
450	460	470	480
TTAGAGCGAT	GGCGGACATA	GTCATTAACC	ACCGTGTGG
490	500	510	520
GAECTACTCAA	GGGCATGGTG	GAATGTACAA	CCGCTATGAT
530	540	550	560
GGAATTCCCTA	TGTCTTGGGA	TGAACATGCT	ATTACATCTT
570	580	590	600
GCACTGGTGG	AAGGGGTAAC	AAAAGCACTG	GAGACAAC
610	620	630	640
TAATGGAGTT	CCAAATATAG	ATCATACACA	ATCCTTTGTT
650	660	670	680
CGGAAAGATC	TCATTGACTG	GATGCCGTGG	CTAAGATCCT
690	700	710	720
CTGTTGGCTT	CCAAGATTTT	CGTTTTGATT	TTGCCAAAGG
730	740	750	760
TTATGCTTCA	AAGTATGTAA	AGGAATATAT	CGAGGGAGCT
770	780	790	800
GAGCCAATAT	TTGCAGTTGG	AGAATACTGG	GACACTTGCA
810	820	830	840
ATTACAAGGG	CAGCAATTG	GATTACAACC	AAGATAGTCA
850	860	870	880
CAGGCAAAGA	ATCATCAATT	GGATTGATGG	CGCGGGACAA
890	900	910	920
CTTCAACTG	CATTCGATT	TACAACAAAA	GCAGTCCTTC

FIGURE 4 CONTINUED

12/38

930	940	950	960
AGGAAGCAGT	CAAAGGAGAA	TCTGGCGTT	TGCGTGACTC
970	980	990	1000
TAAGGGGAAAG	CCCCCAGGAG	TTAGGATT	GTGGCCTTCA
1010	1020	1030	1040
AGGGCTGTC	ATTTTATGAA	TAATCACGAC	ACTGGATCAA
1050	1060	1070	1080
CTCAGGGCGCA	TTGGCCCTTC	CCTTCACGTC	ATGTTATGGA
1090	1100	1110	1120
GGGCTATGCA	TACATTCTTA	CACACCCAGG	GATACCATCA
1130	1140	1150	1160
GT	TTTCTTGT	ACCATTTCTA	CGAATGGGAT
1170	1180	1190	1200
ATGACCAAAT	TGTAAAGCTG	ATTGCTATTG	GGAGGAATCA
1210	1220	1230	1240
AGGCATACAC	AGCCGTTCAT	CTATAAGAAT	TCTTGAGGCA
1250	1260	1270	1280
CAGCCAAACT	TATACGCTGC	AACCATTGAT	GAAAAGGTTA
1290	1300	1310	1320
GCGTGAAGAT	TGGGGACGGA	TCATGGAGCC	CTGCTGGGAA
1330	1340	1350	1360
AGAGTGGACT	CTCGCGACCA	GTGGCCATCG	CTATGCAGTC
1370	1380	1390	1400
TGGCAGAACT	AATCTTACAG	CTATTCCGTT	ACTTAATATA
1410	1420	1430	1440
TTAGTAGAAA	TATATATGTT	TTAAACCCGA	GCACCTACCT
1450	1460	1470	1480
CTAACACTAG	ATCCGCCCTCT	ACAGGGCTTGG	ATGGAGTGAT
1490	1500	1510	1520
GAGTTTTTTT	TTCTGTTC	TTAGACATTG	CAACATGGGA
1530	1540	1550	1560
TGTATGTTT	GTAAATAAAA	GTGTTCTTGA	TCAATGCAAT
1570			
GTAATAAGGG			

FIGURE 5

13/38

SEQUENCE: Nucleotide sequence of a cDNA encoding the large subunit of ADP-glucose pyrophosphorylase from barley seed endosperm (bepl10)
 SEQUENCE TYPE: NUCLEIC ACID
 MOLECULE TYPE: DNA
 ORIGINAL SOURCE: BARLEY
 SEQUENCE LENGTH: 3037
 STRANDEDNESS: DOUBLE
 TOPOLOGY: LINEAR

```

1      ACGACCACCT CGGAACTCAA CGCCTCCACG GACCATCTCT
41     CTCTCTCTCC CTCCTCTCAC CACCAACC ACCACCACCC
81     CCTCTCCCTC CCTGCATTTG ATTCTGTTCAT ATTCTATCCGT
121    CGCTTGGCCG GTGCCACCC CGTCGATCCC TCACCCCCGCC
161    GTCCCCGGCA GTTGCAGGTG GACTGCTAAT GTCATCGATG
201    CAGTTCAAGCA GCGTGTGCC CCTGGAGGGC AAGGCCTGCG
241    TTTCCCCAGT CAGGAGAGAG GGATCGGCCT GCGAGCGCCT
281    CAAGATCGGG GACAGCAGCA GCATCAGGCA CGAGAGAGCG
321    TCCAGGAGGA TGTGCAACGG CGGCGCAGGG GCCCCGCCGCC
361    CACCGGTGCG CAGTGCCTGC TCACCTCCGA CGCCAGCCCG
401    GCCGACACCC TTGTTCTCCG GACGTCCCTC CGGAGGAATT
          ACGCCGATCC GAACGAGGTC GCGGCCGTCG GTCGCGGCCG
          TCATACTCGG CGGCGGCACC GGGACTTCAGC TCTTCCCGCT
          CACAAGCACA AGGGCCACAC CTGCTGTTCC TATTGGAGGA
          TGTTACAGGC TCATCGATAT TCCCATGAGC AACTGCTTCA
601    ACAGTGGCAT CAACAAGATA TTCGTATGA CCCAGTTCAA
          CTCGGCATCT CTCATCGCC ACATTACCG CACCTACCTC
          GGCGGGGGAA TCAATTTCAC TGATGGATCT GTTGAGGTAT
          TGGCCGCGAC ACAAAATGCC GGGGAGGCTG CTGGATGGTT
          CGCGGAAACA GCGGATGCCG TCAGAAAATT TATCTGGGTG
801    CTTGAGGACT ACTATAAGCA TAAATCCATA GAGCACATT
          TGATCTTGTG GGGCGATCAG CTTTATCGCA TGATTACAT
          GGAGCTTGTG CAGAACATG TGGATGACAA TGCTGACATT
          ACTTATCAT GTGCCCTGTG TGGAGAGAGC CGGGCATCTG
          AGTACGGGCT AGTGAAGTTG GACAGTTCA GCGGTGTGAT
1001   CCAGTTTCT GAGAACGCCA AGGGCGACGA TCTGGAAGCG
          ATGAAAGTGG ATACCAAGTT TCTCAATTTC GCCATAGACG
          ACCCTGCTAA ATATCCATAC ATTGCTTCA GGGAGTTTA
          TGCTTCAAG AGAGATGTTG TGCTGAACCT TCTAAAGTCA
          AGATACCGAG AACTACATGA CTTTGGGTCT GAAATCCTCC
1201   CGAGAGCTC GCATGATCAC AATGTACAGG CATATGTCTT
          CACTGACTAC TGGGAGGACA TTGGAACAAT CAGATCCTTC
          TTGATGCGA ACATGGCCCT CTGCGAACAG CCTCCAAAGT
          TTGAATTTC TGATCCAAAA ACCCCCTTCT TCACTTCGCC
          TCGGTACTTA CGGCCAACAA AGTCAGACAA GTGCAGGATC
1401   AAAGAACGCA TCATTCGCA CGGCTGCTTC TTGCGTGAAT
          GCAAAATCGA GCACTCCATC ATCGGCGTTG GTTCACGCCT
          AACTCCGGA AGCGAGCTCA AGAACCGAT GATGATGGGC
          GCGGACTCGT ACGAGACCGA GGACGAGATC TCGAGGCTGA
          TGTCTGAGGG CAAGGTTCCC ATCGGCGTCG GGGAGAACAC
1601   AAAGATCAGC AACTGCATCA TCGACATGAA CGCGAGGATA
  
```

14/38

FIGURE 5 CONTINUED

GGAAGGGACG TGGTCATCTC AAACAAGGAG GGGGTGCAAG
AAGCCGACAG GCCGGAGGAA GGGTACTACA TCAGGTCCGG
GATCGTGGTG ATCCAGAAGA ACGCGACCAT CAAGGACGGC
1801 ACCGTCGTGT AGGGCGTGC GGGTCGGCGC GACGGGGTTC
TGGGACAACC TGTGCGCTGC GTGGTCGTC ATCATCTCT
CAAACCTCCGG GACTGAAGAA GTGATCCGGG GACGGGAGAAC
GTTTGAAGCT TGAATGACTG AGACTGAAAG TGAAGGCGCA
GCAGAGGCAG GCAGCATTAG TAGTAAGTAG TAAGTAAGTA
2001 GCAGTGGAAC AAAGTAATAG TCGTTCGTT TTCCCCCTGTA
ATAAATAAAGA GGCTGTGTGT TGAGGTAAAA AAAAAAA

FIGURE 6

15/38

SEQUENCE: Nucleotide sequence of a cDNA encoding the small subunit of ADP-glucose pyrophosphorylase from barley seed endosperm (beds)
 SEQUENCE TYPE: NUCLEIC ACID
 MOLECULE TYPE: DNA
 ORIGINAL SOURCE: BARLEY
 SEQUENCE LENGTH: 1822
 STRANDEDNESS: DOUBLE
 TOPOLOGY: LINEAR
 COMMENT: The "." at 1569 denotes a purine.

```

1      AAAAGTGAAC TCACACATCA CTCAATATCT ATATCCTTCC
       ATTATATATC CCTCGGTGAT GGATGTACCT TTGGCATCTA
       AAGTTCCCTT GCCTCCCTT TCCAAGCATG AACAAATGCAA
       CGTTTATAGT CATAAGAGCT CATCGAAGCA TGCAAGATCTC
       AATCCCCATG CTATTGATAG TGTTCTCGGT ATCATTTCTTG
201     GAGGTGGTGC AGGGACTAGA TTGTATCCCC TGACGAAGAA
       GCGTGCAAAG CCTGCAGTGC CATTGGGTGC CAACTACAGG
       CTTATTGATA TTCTGTCAAG TAATTGTCTG AACAGCAACA
       TATCAAAGAT CTATGTGCTT ACACAGTTCA ACTCAGCTTC
       TCTTAATCGT CATCTCTCAC GAGCCTATGG GAGCAACATT
       GGAGGTTACA AGAATGAAGG ATTTGTTGAA GTCCCTTGCTG
       CACAGCAGAG CCCAGATAAC CCTGACTGGT TCCAGGGTAC
       TGCAGATGCT GTAAGGCAGT ACTTGTGGCT ATTGAGGGAG
       CATAATGTTA TGGAGTATCT AATTCTTGCT GGAGATCACC
       TGTACCGAAT GGACTATGAA AAGTTTATTG AGGCACACAG
       AGAAACGGAT GCTGATATTA CTGTTGCTGC CTTGCCCATG
       GATGAGGAAC GTGCAACTGC ATTTGGCCTT ATGAAAATCG
       ATGAAGAAGG GAGGATAATT GAATTGCGAG AGAAACCAAA
       AGGAGAACAG TTGAAAGCTA TGATGGTTGA TACGACCATA
       CTTGGCCTTG AAGATGCGAG GGCAAAGGAA ATGCCTTATA
       801     TTGCTAGCAT GGGTATCTAT GTTATTAGCA AACATGTGAT
       GCTTCAGCTT CTCCGTGAGC AATTTCCTGG AGCTAATGAC
       TTCGGAAGTG AAGTTATCCC TGGTGCAACT AGCACTGGCA
       TGAGGGTACA AGCATAACCTA TACGACGGTT ACTGGGAAGA
       TATTGGTACA ATTGAGGCAT TCTATAATGC AAATTTGGGA
       1001    ATTACCAAAAA AACCAATACC TGATTTCACT TTCTATGACC
       GTTCTGCTCC CATTACACA CAACCTCGAC ACTTGCCTCC
       TTCAAAGGTT CTTGATGCTG ATGTGACAGA CAGTGTAAATT
       GGTGAAGGAT GTGTTATTAA AAACGTCAAG ATACACCATT
       CAGTAGTTGG ACTCCGTTCC TGCAATATCTG AAGGTGCAAT
       1201    AATAGAGGAC ACGTTGCTAA TGGGTGCGGA CTACTATGAG
       ACTGAAGCTG ATAAGAAACT CCTTGCTGAA AAAGGTGGCA
       TTCCCATTGG TATTGGAAAG AATTCACACA TCAAAAGAGC
       AATCATTGAC AAGAATGCTC GTATTGGAGA TAACGTGATG
       ATAATCAATG TTGACAATGT TCAAGAAGCG GCGAGGGAGA
       1401    CAGATGGATA CTTCATCAAA AGTGGCATCG TAACTGTGAT
       CAAGGATGCT TTACTCCCTA GTGGAACAGT CATATGAAGC
       AGATGTGAAA TGTATGCCAA AAGACAGGGC TACTTGCCTC
       AGTCTGGAAAT CAACCAACAA GGCGCGGAAG GAGATCATAA
       ATAACCAAG GAGTGCCTG CGAGTCACCTT CTACACCCCTT
       1601    TTCCCCCCTT GATGTATTAG GAACTGTGAT GTACAAGCAA
  
```

FIGURE 6 CONTINUED

16/38

CTGTGATGCA CTTACGCGAA GTGCCCTGG ATTCAGCTTT
CTCTTTGCTT GTAACTGGTT TCCAGCAGAC CATGCTATTT
GTTGTATGGT TCGTGCAAAA CCTTGCGATG CTTTATATAT
GCTTTATATA TAAACAAGAT GAATCCCCGC GCGTTGCTGC
2001 GGCACAAAAA AAAAAAAA AA

17/38

FIGURE 7

α -GLUCAN LYASE CODING SEQUENCE
 SEQUENCE TYPE: NUCLEIC ACID
 MOLECULE TYPE: DNA (GENOMIC)
 ORIGINAL SOURCE: FUNGALLY INFECTED ALGAE
 SEQUENCE LENGTH 3267 BP
 STRANDEDNESS: DOUBLE
 SEQUENCE:

	10	20	30	40	50	60
1	ATGTTTCAA	CCCTTGCGT	TGTGGCACCT	AGTGCCTGG	GAGCCAGTAC	CTTGTAGGG
61	GCGGAGGTCA	GGTCATAATG	TCGTATCCAT	TCCGCTTTC	CAGCTGTGCA	CACAGCTACT
121	CGCAAAACCA	ATCGCCTCAA	TGTATCCATG	ACCGCATTGT	CCGACAAACA	AACGGCTACT
181	GCGGGTAGTA	CAGACAATCC	GGACGGTATC	GACTACAAGA	CCTACGATTA	CGTCGGAGTA
241	TGGGGTTTCA	GCCCCCTCTC	CAACACGAAC	TGGTTTGTG	CCGGCTCTC	TACCCCCGGT
301	GGCATCACTG	ATTGGACGGC	TACAATGAAT	GTCAACTTCG	ACCGTATCGA	CAATCGCTC
361	ATCACTGTCC	AGCATCCCGT	TCAGGTTCA	GTCACGTCAT	ACAACAACAA	CAGCTACAGG
421	GTTCGCTTCA	ACCCCTGATGG	CCCTATTGCGT	GATGTGACTC	GTGGGCTAT	CCTCAAGCAG
481	CAACTAGATT	GGATTCGAAC	GCAGGAGCTG	TCAGAGGGAT	GTGATCCCCG	AATGACTTTC
541	ACATCAGAAG	GTTCCTTGAC	TTTGAGACC	AAGGATCTAA	GGCTCATCAT	CTACGGAAAT
601	TTCAAGACCA	GAGTTACGAG	AAAGTCTGAC	GGCAAGGTCA	TCATGGAAAA	TGATGAAGTT
661	GGAACATGCAT	CGTCCGGGAA	CAAGTGCCGG	GGATTGATGT	TCGTTGATAG	ATTATACGGT
721	AACGCTATCG	CTTCCGTCAA	CAAGAACCTC	CGCAACGACG	CGTCAGCAAGCA	GGAGGGATT
781	TATGGTCAG	GTGAAGTCAA	CTGTAAGTAC	CAGGACACCT	ACATCTTAGA	ACGCAC TGGA
841	ATCGCCATGA	CAAATTACAA	CTACGATAAC	TTGAACATA	ACCACTGGGA	CTTAGACCT
901	CCGCATCATG	ATGGTGCCCT	CAACCCAGAC	TATTATATTC	CAATGTACTA	CGCAGCACCT
961	TGGTTGATCG	TTAATGGATG	CGCCGGTACT	TCGGAGCAGT	ACTCGTATGG	ATGGTTCATG
1021	GACAATGTCT	CTCAATCTTA	CATGAATACT	GGAGATACTA	CCTGGAAATT	TGGACAAGAG
1081	GACCTGGCAT	ACATGGGCGC	GCAGTATGGA	CCATTTGACC	AACATTTTGT	TTACGGTGCT
1141	GGGGGTGGGA	TGGAATGTGT	GGTCACAGCG	TTCTCTCTTC	TACAAGGCAA	GGAGTTCGAG
1201	AACCAAGTTC	TCAACAAACG	TTCAGTAATG	CCTCCGAAAT	ACGTCCTTGG	TTCTTCCAG
1261	GGTGTTCG	GGACTTCTTC	CTTGTGAGA	GCGCATATGC	CAGCAGGTGA	GAACAACATC
1321	TCAGTCGAAG	AAATTGTAGA	AGGTTATCAA	AACAACAATT	TCCCCTTCGA	GGGGCTCGCT
1381	GTGGACGTGG	ATATGCAAGA	CAACTGCGG	GTGTTCACCA	CGAAGGGCGA	ATTTTGGACC
1441	GCAAACAGGG	TGGGTACTGG	CGGGGATCCA	AAACAACCGAT	CGGTTTTGA	ATGGGCACAT
1501	GACAAAGGCC	TTGTTTGTCA	GACAAATATA	ACTTGCTTCC	TGAGGAATGA	TAACGAGGGG
1561	CAAGACTACG	AGGTCAATCA	GACGTTAAGG	GAGAGGCACT	TGTACACGAA	GAACGACTCC
1621	CTGACGGGTA	CGGATTTGG	AATGACCGAC	GACGGCCCCA	GCGATTCGTA	CATCGGTCT
1681	CTGGACTATG	GGGGTGGAGT	AGAATGTAT	GCACCTTTCC	CAGACTGGGG	ACGGCCTGAC
1741	GTGGCCGAAT	GGTGGGGAAA	TAACTATAAG	AAACTGTTCA	GCATTGGTCT	CGACTTCGTC
1801	TGGCAAGACA	TGACTGTTCC	AGCAATGATG	CGCACA AAA	TTGGCGATGA	CATCAATGTG
1861	AAACCGGATG	GGAAATTGGCC	GAATGCGGAC	GATCCGTCA	ATGGACAATA	CAACTGGAAG
1921	ACGTACCATC	CCCAAGTGC	TGTAACGTAT	ATGCGTTATG	AGAATCATGG	TCGGGAACCG
1981	ATGGTCACTC	AACGCAACAT	TCATGCGTAT	ACACTGTGCG	AGTCTACTAG	GAAGGAAGGG
2041	ATCGTGGAAA	ACGCAGACAC	TCTAACGAAG	TTCCGCCGTA	GCTACATTAT	CAGTCGTGGT
2101	GGTTACATTG	GTAACCAGCA	TTTCGGGGT	ATGTGGGTGG	GAGACAACTC	TACTACATCA
2161	AACTACATCC	AAATGATGAT	TGCCAACAT	ATTAACATGA	ATATGCTT	CTTGCCTCTC
2221	GTCGGCTCCG	ACATTGGAGG	ATTACCTCA	TACGACAATG	AGAATCAGCG	AACGCCGTGT
2281	ACCGGGGACT	TGATGGTGA	GTATGCGAG	GCGGGCTGCC	TGTTGCCGTG	GTTAGGAAC
2341	CACTATGATA	GGTGGATCGA	GTCCAAGGAC	CACGGAAAGG	ACTACCAAGGA	GCTGTACATG
2401	TATCCGAATG	AAATGGATAC	GTTGAGGAAG	TTCGTTGAAT	TCCGTTATCG	CTGGCAGGAA
2461	GTGTTGTACA	CGGCCATGTA	CCAGAATGCG	GCTTTCGGAA	AGCCGATTAT	CAAGGCTGCT
2521	TCGATGTACA	ATAACGACTC	AAACGTTCGC	AGGGCGCAGA	ACGATCA	CCTTCTGGT
2581	GGACATGATG	GATATCGCAT	TCTGTGCGCG	CCTGTTGTGT	GGGAGAATT	GACCGAACGC

FIGURE 7 CONTINUED

2641 GAATTGTACT TGCCCGTGCT GACCCAATGG TACAAATTCTG GTCCCGACTT TGACACCAAG
2701 CCTCTGGAAG GAGCGATGAA CGGAGGGGAC CGAATTACA ACTACCCCTGT ACCGCAGAAGT
2761 GAATCACCAA TCTTCGTGAG AGAAGGTGCG ATTCTCCTA CCCGCTACAC GTTGAACGGT
2821 GAAAACAAAT CATTGAACAC GTACACGGAC GAAGATCCGT TGTTGTTGA AGTATTCCCC
2881 CTCGGAAACA ACCGTGCCGA CGGTATGTGT TATCTTGATG ATGGCGGTGT GACCACCAAT
2941 GCTGAAGACA ATGGCAAGTT CTCTGTGTC AAGGTGGCAG CGGAGCAGGA TTGGTGGTACG
3001 GAGACGATAA CGTTTACGAA TGATTGCTAT GAGTACGTT TCGGTGGACC GTTCTACGTT
3061 CGAGTGCAGCG GCGCTCAGTC GCGTCGAAC ATCCACGTGT CTCTGGAGC GGGTTCTCAG
3121 GACATGAAGG TGAGCTCTGC CACTTCCAGG GCTGCCTGT TCAATGACGG GGAGAACGGT
3181 GATTTCTGGG TGACCCAGGA GACAGATTCT CTGTGGCTGA AGTTGCCAA CGTTGTTCTC
3241 CGGGACGCTG TGATCACAAAT TACCTAA

FIGURE 8

19/38

α -GLUCAN LYASE CODING SEQUENCE
 SEQUENCE TYPE: NUCLEIC ACID
 MOLECULE TYPE: DNA (GENOMIC)
 ORIGINAL SOURCE: FUNGALLY INFECTED ALGAE
 SEQUENCE LENGTH: 3276 BP
 STRANDEDNESS: DOUBLE
 SEQUENCE.

	10	20	30	40	50	60
1	ATGTATCCAA	CCCTCACCTT	CGTGGCGCCT	AGTGCGCTAG	GGGCCAGAAC	TTCACGTGT
61	GTGGGCATT	TTAGGTCACA	CATTCTTATT	CATTCCGGTTG	TTCCAGCGGT	GCGTCTAGCT
121	GTGCGCAAAA	GCAACCGCCT	CAATGTATCC	ATGTCCGCTT	TGTTGACCAA	ACCGACTGCT
181	GTTACTGGAG	GGAAGGACAA	CCCGGACAAT	ATCAATTACA	CCACTTATGA	CTACGTCCT
241	GTGTGGCGCT	TCGACCCCC	CAGCAATAAG	AACCTGGTTTG	CTGCCGGATC	TTCCACTCCC
301	GGCGATATTG	ACGACTGGAC	GGCGACAATG	AATGTGAAC	TCGACCGTAT	CGACAATCCA
361	TCCTTCACTC	TCGAGAAACC	GGTTCAGGTT	CAGGTACAGT	CATACAAGAA	CAATTGTTTC
421	AGGGTTCGCT	TCAACCCCTGA	TGGTCTTATT	CGCGATGTGG	ATCGTGGGCC	TATCCTCCAG
481	CAGCAACTAA	ATTGGATCCG	GAAGCAGGAG	CAGTCGAAGG	GGTTTGATCC	TAAGATGGC
541	TTCACAAAAG	AAGGTTTCTT	GAAATTGAG	ACCAAGGATC	TGAACGTTAT	CATATATGGC
601	AATTTTAAGA	CTAGAGTTAC	GAGGAAGAGG	GATGGAAAAG	GGATCATGGA	GAATAATGAA
661	GTGCCGGCAG	GATCGTTAGG	GAACAAGTGC	CGGGGATTGA	TGTTTGTGA	CAGGTTGTAC
721	GGCACTGCCA	TCGCTTCCGT	TAATGAAAAT	TACCGCAACG	ATCCCGACAG	GAAAGAGGGG
781	TTCTATGGTG	CAGGAGAAGT	AAACTGCGAG	TTTTGGACT	CCGAACAAAAA	CAGGAACAAG
841	TACATCTTAG	AACGAACCTGG	AATGCCATG	ACAAATTACA	ATTATGACAA	CTATAACTAC
901	AACCAGTCAG	ATCTTATTGC	TCCAGGATAT	CCTTCCGACC	CGAACTTCTA	CATTCCCATG
961	TATTTTGCA	CACCTTGGGT	AGTTGTTAAG	GGATGCACTG	GCAACAGCGA	TGAACAGTAC
1021	TCGTACGGAT	GGTTTATGGA	TAATGTCCTC	CAAACATTACA	TGAATACTGG	TGGTACTTCC
1081	TGGAACTGTG	GAGAGGGAGA	CTTGGCATAAC	ATGGGAGCAC	AGTGCCTGTC	ATTTGACCAA
1141	CATTTTGTGT	ATGGTGTATGG	AGATGGTCTT	GAGGATGTTG	TCCAAGCGTT	CTCTCTTCTG
1201	CAAGGCAAAG	AGTTTGAGAA	CCAAGTTCTG	AACAAACGTG	CCGTAATGCC	TCCGAAATAT
1261	GTGTTTGGTT	ACTTTTCAAGGG	AGTCTTGGG	ATTGCTTCT	TGTTGAGAGA	GCAAAGACCA
1321	GAGGGTGGTA	ATAAACATCTC	TGTTCAAGAG	ATTGTCGAAG	GTTACCAAAG	CAATAACTTC
1381	CCTTTAGAGG	GGTTAGCCGT	AGATGTGGAT	ATGCAACAAAG	ATTGCGCGT	GTTACCAACAG
1441	AAGATTGAAT	TTGGACGGC	AAATAAGGTA	GGCACCGGGG	GAGACTCGAA	TAACAAGTCG
1501	GTGTTTGAAT	GGGCACATGA	CAAAGGCCCTT	GTATGTCAGA	CGAATGTTAC	TTGCTTCTTG
1561	AGAAACGACA	ACGGCGGGGGC	AGATTACGAA	GTCAATCAGA	CATTGAGGGA	GAAGGGTTTG
1621	TACACGAAGA	ATGACTCACT	GACGAACACT	AACTTCGAA	CTACCAACGA	CGGGCCGAGC
1681	GATGCGTACA	TTGGACATCT	GGACTATGGT	GGCGGGAGGGA	ATTGTGATGC	ACTTTTCCCA
1741	GAATGGGGTC	GACCAGGGTGT	GGCTGAATGG	TGGGGTGATA	ACTACAGCAA	GCTCTTCAA
1801	ATTGGTCTGG	ATTTCGTCTG	GCAAGACATG	ACAGTTCCAG	CTATGATGCC	ACACAAAGTT
1861	GGCGACCGAG	TCGATACGAG	ATCACCTTAC	GGCTGGCCGA	ATGAGAAC	TCCTTCGAAC
1921	GGACGATACA	ATTGGAAATC	TTACCATCCA	CAAGTTCTG	TAACTGATAT	GCGATATGAG
1981	AATCATGGAA	GGGAACCGAT	GTTCACCTAA	CGCAATATGC	ATGCGTACAC	ACTCTGTGAA
2041	TCTACGAGGA	AGGAAGGGAT	TGTTGCAAAT	GCAGACACTC	TAACGAAGTT	CCGCGCGAGT
2101	TATATTATCA	GTCGTGGAGG	TTACATTGGC	AACCAGCATT	TTGGAGGAAT	GTGGGTTGG
2161	GACAACCTT	CCTCCCAAAG	ATACCTCCAA	ATGATGATCG	CGAACATCGT	CAACATGAAC
2221	ATGTCTTGC	TTCCACTAGT	TGGGTCCGAC	ATTGGAGGTT	TTACTTCGTA	TGATGGACGA
2281	AACGTGTGTC	CCGGGGATCT	AATGGTAAGA	TTCGTGCAAG	CGGGTTGCTT	ACTACCGTGG
2341	TTCAGAAACC	ACTATGGTAG	GTTGGTCGAG	GGCAAGCAAG	AGGGAAAATA	CTATCAAGAA
2401	CTGTACATGT	ACAAGGACGA	GATGGCTACA	TTGAGAAAAT	TCATTGAATT	CCGTTACCGC
2461	TGGCAGGAGG	TGTTGTACAC	TGCTATGTAC	CAGAATGCAGG	CTTTCGGGAA	ACCGATTATC
2521	AAGGCAGCTT	CCATGTACGA	CAACGACAGA	ACGTTCGCG	GCGCACAGGA	TGACCAACTTC
2581	CTTCTCGGCG	GACACGATGG	ATATCGTATT	TTGTGTGCAC	CTGTTGTGTG	GGAGAATACA

FIGURE 3 CONTINUED

2641 ACCAGTCGCG ATCTGTACTT GCCTGTGCTG ACCAAATGGT ACAAAATTGGG CCCTGACTAT
2701 GACACCAAGC GCCTGGATTG TGCGTTGGAT GGAGGGCAGA TGATTAAGAA CTATTCTGTG
2761 CCACAAAGCG ACTCTCCGAT ATTGTGAGG GAAGGAGCTA TTCTCCCTAC CCGCTACACG
2821 TTGGACGGTT CGAACAAAGTC AATGAACACG TACACAGACA AAGACCCGTT GGTGTTTGAG
2881 GTATTCCTC TTGGAAACAA CCGTGCCGAC GGTATGTGTT ATCTTGTGA TGGCGGTATT
2941 ACTACAGATG CTGAGGACCA TGGCAAATTG TGTGTTATCA ATGTCGAAGC C-TACGGAAA
3001 GGTGTTACGA CGACGATCAA GTTTGCSTAT GACACTTATC AATACTGATTT TGTGGTCCA
3061 TTCTACGTTG GAATCCGTAATCTTACGACT GCATCAAAAAA TTACGTGTC TTGTGGAGCG
3121 GGTGAAGAGG ACATGACACC GACCTCTGCG AACCTGAGGG CAGCTTGTGTT CAGTGATGGA
3181 GGTGTTGGAG AATACTGGGC TGACAATGAT ACGTCTTCTC TGTGGATGAA G-TGCCAAAC
3241 CTGGTTCTGC AAGACGCTGT GATTACCAATT ACGTAG

FIGURE 9

21/38

α -GLUCAN LYASE CODING SEQUENCE
 SEQUENCE TYPE: NUCLEIC ACID
 MOLECULE TYPE: DNA (GENOMIC)
 ORIGINAL SOURCE: FUNGUS
 SEQUENCE LENGTH: 3201 BP
 STRANDEDNESS: DOUBLE
 SEQUENCE:

10	20	30	40	50	60
ATGGCAGGAT	TTTCTGATCT	TCTCAACTTT	TGCAAAGCAG	AAGACTACTA	CAGTGTTCG
70	80	90	100	110	120
CTAGACTGGA	AGGGCCCTCA	AAAAATCATT	GGAGTAGACA	CTACTCCTCC	AAAGAGCACC
130	140	150	160	170	180
AAGTTCCCCA	AAAACCTGGCA	TGGAGTGAAC	TTGAGATTG	ATGATGGGAC	TTTAGGTG
190	200	210	220	230	240
GTTCAAGTTCA	TTAGGCCGTG	CGTTGGAGG	GTTAGATACTG	ACCCCTGGTT	CAAGACCTCT
250	260	270	280	290	300
GACGAGTATG	GTGATGAGAA	TACGAGGACA	ATTGTGCAAG	ATTATATGAG	TACTCTGAGT
310	320	330	340	350	360
AATAAAATTGG	ATACTTATAG	AGGTCTTACG	TGGGAAACCA	AGTGTGAGGA	TTCGGGAGAT
370	380	390	400	410	420
TTCTTTACCT	TCTCATCCAA	GGTCACCGCC	GTTGAAAAAT	CCGAGCGGAC	CCGCAACAAG
430	440	450	460	470	480
GTCGGCGATG	GCCTCAGAAC	TCACCTATGG	AAAAGCCCTT	TCCGCATCCA	AGTAGTGCGC
490	500	510	520	530	540
ACCTTGACCC	CTTTGAAGGA	TCCCTACCCC	ATTCCAAATG	TAGCCGCAGC	CGAAGCCCGT
550	560	570	580	590	600
GTGTCCGACA	AGGTCGTTTG	GCAAACGTCT	CCCAAGACAT	TCAGAAAGAA	CCTGCATCCG
610	620	630	640	650	660
CAACACAAGA	TGCTAAAGGA	TACAGTTCTT	GACATTGTCA	AACCTGGACA	TGGCGAGTAT
670	680	690	700	710	720
GTGGGGTGGG	GAGAGATGGG	AGGTATCCAG	TTTATGAAGG	AGCCAACATT	CATGAACTAT
730	740	750	760	770	780
TTTAACCTTCG	ACAATATGCA	ATACCAGCAA	GTCTATGCC	AAGGTGCTCT	CGATTCTCGC
790	800	810	820	830	840
GAGCCACTGT	ACCACTCGGA	TCCCTTCTAT	CTTGATGTGA	ACTCCAACCC	GGAGCACAAAG
850	860	870	880	890	900
AATATCACCG	CAACCTTTAT	CGATAACTAC	TCTCAAATTG	CCATCGACTT	TGGAAAGACC
910	920	930	940	950	960
AACTCAGGCT	ACATCAAGCT	GGGAACCAGG	TATGGTGGTA	TCGATTGTTA	CGGTATCAGT
970	980	990	1000	1010	1020
GCGGATACGG	TCCCGGAAAT	TGTACGACTT	TATACAGGTC	TTGTTGGACG	TTCAAAGTTG
1030	1040	1050	1060	1070	1080
AAGCCCAGAT	ATATTCTCGG	GGCCCATCAA	GCCTGTTATG	GATACCAACA	GGAAAGTGA
1090	1100	1110	1120	1130	1140
TTGTATTCTG	TGGTCCAGCA	GTACCGTGAC	TGTAAATTTC	CACTTGACGG	GATTCACTGTC
1150	1160	1170	1180	1190	1200
GATGTCGATG	TTCAGGACGG	CTTCAGAACT	TTCACCAACCA	ACCCACACAC	TTTCCCTAAC
1210	1220	1230	1240	1250	1260
CCCAAAGAGA	TGTTTACTAA	CTTGAGGAAT	AATGGAATCA	AGTGCTCCAC	CAATATCACT
1270	1280	1290	1300	1310	1320
CCTGTTATCA	GCATTAACAA	CAGAGAGGGT	GGATACAGTA	CCCTCCTTGA	GGGAGTTGAC

FIGURE 9 CONTINUED

1330	1340	1350	1360	1370	1380
AAAAAAATACT	TTATCATGGA	CGACAGATAT	ACCGAGGGAA	CAAGTGGGAA	TGCGAAGGAT
1390	1400	1410	1420	1430	1440
GTTCGGTACA	TGTACTACGG	TGGTGGTAAT	AAGGTTGAGG	TCGATCCTAA	TGATGTTAAT
1450	1460	1470	1480	1490	1500
GGTCGGCCAG	ACTTTAAAGA	CAACTATGAC	TTCCCCGCGA	ACTTCAACAG	CAAACAATAC
1510	1520	1530	1540	1550	1560
CCCTATCATG	GTGGTGAG	CTACGGTTAT	GGGAACGGTA	GTGCAGGTTT	TTACCCGGAC
1570	1580	1590	1600	1610	1620
CTCAACAGAA	AGGAGGGTCG	TATCTGGTGG	GGAATGCAGT	ACAAGTATCT	CTTCGATATG
1630	1640	1650	1660	1670	1680
GGACTGGAAT	TTGTGTGGCA	AGACATGACT	ACCCCAGCAA	TCCACACATC	ATATGGAGAC
1690	1700	1710	1720	1730	1740
ATGAAAGGGT	TGCCAACCG	TCTACTCGTC	ACCTCAgACT	CCTCACCAA	TGCCTCTGAG
1750	1760	1770	1780	1790	1800
AAAAAGCTG	CAATTGAAAC	TTGGGCTCTC	TACTCCTACA	ATCTCCACAA	AGCAACTTGG
1810	1820	1830	1840	1850	1860
CATGGTCTTA	GTCGTCTCGA	ATCTCGTAAG	AACAAACGAA	ACTTCATCCT	CGGGCGTGGA
1870	1880	1890	1900	1910	1920
AGTTATGCCG	GAGCCTATCG	TTTTGCTGGT	CTCTGGACTG	GGGATAATGC	AAGTAACTGG
1930	1940	1950	1960	1970	1980
GAATTCTGGA	AGATATCGGT	CTCTCAAGTT	CTTTCTCTGG	GCCTCAATGG	TGTGTGCATC
1990	2000	2010	2020	2030	2040
GCAGGGGTCTG	ATACGGGTGG	TTTTGAACCC	TACCGTGATG	CAAATGGGGT	CGAGGAGAAA
2050	2060	2070	2080	2090	2100
TACTGTAGCC	CAGAGCTACT	CATCAGGTGG	TATACTGGTT	CATTCCCTCTT	GCCGTGGCTC
2110	2120	2130	2140	2150	2160
AGGAACCATT	ATGTCAAAAA	GGACAGGAAA	TGGTTCCAGG	AACCATACTC	GTACCCCAAG
2170	2180	2190	2200	2210	2220
CATCTGAAA	CCCATCCAGA	ACTCGCAGAC	CAAGCATGGC	TCTATAAATC	CGTTTGGAG
2230	2240	2250	2260	2270	2280
ATCTGTAGGT	ACTATGTGGA	GCTTAGATAC	TCCCTCATCC	AACTACTTTA	CGACTGCATG
2290	2300	2310	2320	2330	2340
TTTCAAAACG	TaGTGACGG	TATGCCAATC	ACCAGATCTA	TGCTCTTGAC	CGATACTGAG
2350	2360	2370	2380	2390	2400
GATACCACCT	TCTTCAACGA	GAGCCAAAAG	TTCCTCGACA	ACCAATATAT	GGCTGGTGAC
2410	2420	2430	2440	2450	2460
GACATTCTTG	TTGCACCCAT	CCTCCACAGT	CGCAAAGAAA	TTCCAGGCGA	AAACAGAGAT
2470	2480	2490	2500	2510	2520
GTCTATCTCC	CTCTTTACCA	CACCTGGTAC	CCCTCAAATT	TGAGACCATG	GGACGATCAA
2530	2540	2550	2560	2570	2580
GGAGTCGCTT	TGGGGAAATCC	TGTGAAAGGT	GGTAGTGTCA	TCAATTATAC	TGCTAGGATT
2590	2600	2610	2620	2630	2640
GTTGCACCCG	AGGATTATAA	TCTCTTCCAC	AGCGTGGTAC	CAGTCTACGT	TAGAGAGGGT
2650	2660	2670	2680	2690	2700
GCCATCATCC	CGCAAATCGA	AGTACGCCAA	TGGACTGGCC	AGGGGGGAGC	CAACCGCCTC
2710	2720	2730	2740	2750	2760
AAGTTCAACA	TCTACCCCTGG	AAAGGATAAG	GAGTACTGTA	CCTATCTTGA	TGATGGTGTT
2770	2780	2790	2800	2810	2820
AGCCGTGATA	GTGCGCCCGA	AGACCTCCCA	CAGTACAAAG	AGACCCACGA	ACAGTCGAAG
2830	2840	2850	2860	2870	2880
GTTGAAGGCG	CGGAAATCGC	AAAGCAGATT	GGAAAGAAGA	CGGGTTACAA	CATCTCAGGA
2890	2900	2910	2920	2930	2940

FIGURE 9 CONTINUED

23/38

ACCGACCCAG AAGCAAAGGG TTATCACCGC AAAGTTGCTG TCACACAAAC GTCAAAAGAC
2950 2960 2970 2980 2990 3000
AAGACGCCGT A CTGTCACTAT TGAGCCAAAA CACAATGGAT ACGACCCCTTC CAAAGAGGTG
3010 3020 3030 3040 3050 3060
GGTGATTATT ATACCACATCAT TCTTTGGTAC GCACCGAGGT TCGATGGCAG CATCGTCGAT
3070 3080 3090 3100 3110 3120
GTGAGCAAGA CGACTGTGAA TGTGAGGGT GGGGTGGAGC ACCAAGTTA TAAGAACTCC
3130 3140 3150 3160 3170 3180
GATTTACATA CGGTTGTAT CGACGTGAAG GAGGTGATCG GTACCCACAAA GAGCGTCAAG
3190 3200
ATCACATGTA CTGCCGCTTA A

FIGURE 10

24/38

α -GLUCAN LYASE CODING SEQUENCE
 SEQUENCE TYPE: NUCLEIC ACID
 MOLECULE TYPE: DNA (GENOMIC)
 ORIGINAL SOURCE: FUNGUS
 SEQUENCE LENGTH 3213 BP
 STRANDEDNESS: DOUBLE
 SEQUENCE:

10	20	30	40	50	60
ATGGCAGGAT	TATCCGACCC	TCTCAATTTC	TGCAAAGCAG	AGGACTACTA	CGCTGCTGCC
70	80	90	100	110	120
AAAGGCTGGA	GTGGCCCTCA	GAAGATCATT	CGCTATGACC	AGACCCCTCC	TCAGGGTACA
130	140	150	160	170	180
AAAGATCGA	AAAGCTGGCA	TGCGGTAAAC	CTTCCTTTCG	ATGACGGGAC	TATGTGTGTA
190	200	210	220	230	240
GTGCAATTTCG	TCAGACCCCTG	TGTTTGAGG	GTTAGATAATG	ACCCCCAGTGT	CAAGACTTCT
250	260	270	280	290	300
GATGAGTACG	GCGATGAGAA	TACGAGGACT	ATTGTACAAG	ACTACATGAC	TACTCTGGTT
310	320	330	340	350	360
GGAAACTTGG	ACATTTTCAG	AGGTCTTACG	TGGGTTTCTA	CGTTGGAGGA	TTCGGGCGAG
370	380	390	400	410	420
TAATCACACCT	TCAAGTCCGA	AGTCACTGCC	GTGGACGAAA	CCGAACGGAC	TCGAAACAAG
430	440	450	460	470	480
GTCGGCGACG	GCCTCAAGAT	TTACCTATGG	AAAAATCCCT	TTTCGCATCCA	GGTAGTGCCT
490	500	510	520	530	540
CTCTTGACCC	CCCTGGTGGA	CCCTTTCCCC	ATTCCCAACG	TAGCCAATGC	CACAGCCCGT
550	560	570	580	590	600
GTGGCCGACA	AGGTTGTTTG	GCAGACGTCC	CCGAAGACGT	TCAGGAAAAAA	CTTGCATCCG
610	620	630	640	650	660
CAGCATAAGA	TGTTGAAGGA	TACAGTTCTT	GATATTATCA	AGCCGGGGCA	CGGAGAGTAT
670	680	690	700	710	720
GTGGGTTGGG	GAGAGATGGG	AGGCATCGAG	TTTATGAAGG	AGCCAACATT	CATGAATTAT
730	740	750	760	770	780
TTCAACTTG	ACAATATGCA	ATATCAGCAG	GTCTATGCAC	AAGGCCTCT	TGATAGTCGT
790	800	810	820	830	840
GAGCCGTTGT	ATCACTCTGA	TCCCTTCTAT	CTCGACGTGA	ACTCCAACCC	AGAGCACAAG
850	860	870	880	890	900
AACATTACGG	CAACCTTAT	CGATAACTAC	TCTCAGATTG	CCATCGACTT	TGGGAAGACC
910	920	930	940	950	960
AACTCAGGCT	ACATCAAGCT	GGGTACCAGG	TATGGCGGTA	TCGATTGTTA	CGGTATCAGC
970	980	990	1000	1010	1020
GCGGATACGG	TCCC GGAGAT	TGTGCGACTT	TATACTGGAC	TTGTTGGCG	TTCGAAGTTG
1030	1040	1050	1060	1070	1080
AAGCCCAGGT	ATATTCTCGG	AGCCCACCAA	GCTTGTTATG	GATACCAGCA	GGAAAGTGAC
1090	1100	1110	1120	1130	1140
TTGCATGCTG	TTGTTCA GCA	GTACCGTGAC	ACCAAGTTTC	CGCTTGATGG	GTTGCATGTC
1150	1160	1170	1180	1190	1200
GATGTCGACT	TTCAGGACAA	TTTCAGAACG	TTTACCACTA	ACCCGATTAC	GTTCCCTAAT
1210	1220	1230	1240	1250	1260
CCCAAAGAAA	TGTTTACCAA	TCTAAGGAAC	AATGGAATCA	AGTGTCCAC	CAACATCACC
1270	1280	1290	1300	1310	1320

25/38

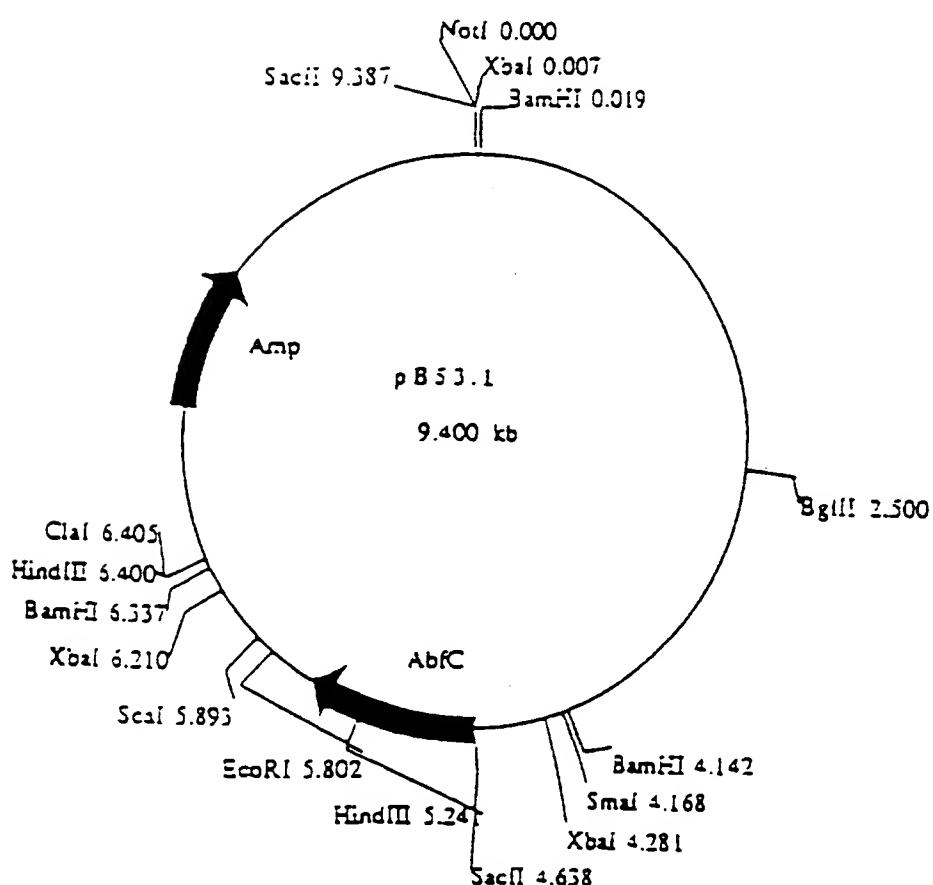
FIGURE 10 CONTINUED

CCTGTTATCA	GTATCAGAGA	TGCCCCGAAT	GGGTACAGTA	CCCTCAATGA	GGGATATGAT
1330	1340	1350	1360	1370	1380
AAAAAGTACT	TGATCATGGA	TGACAGATAT	ACCGAGGGAA	CAAGTGGGAA	CCCGCAAAAT
1390	1400	1410	1420	1430	1440
GTTCGATACT	CTTTTACGG	GGGTGGGAAC	CCGGTTGAGG	TTAACCCCTAA	TGATGTG
1450	1460	1470	1480	1490	1500
GCTCGGGCAG	ACTTTGGAGA	CAATTATGAC	TTCCCTACGA	ACTTCAACTG	CAAAGACTAC
1510	1520	1530	1540	1550	1560
CCCTATATG	GTGGTGTGAG	TTACGGATAT	GGGAATGGCA	CTCCAGGTTA	CTACCGTAC
1570	1580	1590	1600	1610	1620
CTTAACAGAG	AGGAGGTTCG	TATCTGGTGG	GGATTGCACT	ACGAGTATCT	CTTCATATG
1630	1640	1650	1660	1670	1680
GGACTAGAGT	TTGTATGGCA	AGATATGACA	ACCCCAGCGA	TCCATTCACTC	ATATGGAGAC
1690	1700	1710	1720	1730	1740
ATGAAAGGGT	TGCCCCACCG	TCTGCTCGTC	ACCGCCGACT	CAGTTACCAA	TGCCCTGTGAG
1750	1760	1770	1780	1790	1800
AAAAAGCTG	CAATTGAAAG	TTGGGCTCTT	TACTCCTACA	ACCTCCATAA	AGCAACCTTC
1810	1820	1830	1840	1850	1860
CACGGTCTTG	GTCGTCTTGA	GTCTCGTAAG	AACAAACGTA	ACTTCATCCT	CGGACGTGGT
1870	1880	1890	1900	1910	1920
AGTTACGCCG	GTGCCTATCG	TTTGCTGGT	CTCTGGACTG	GAGATAACGC	AAGTACGTGG
1930	1940	1950	1960	1970	1980
GAATTCTGGA	AGATTTGGT	CTCCCCAAGTT	CTTTCTCTAG	GTCTCAATGG	TGTGTGTATA
1990	2000	2010	2020	2030	2040
GCGGGGTCTG	ATACGGGTGG	TTTGAGCCC	GCACGTACTG	AGATTGGGAA	GGAGAAATAT
2050	2060	2070	2080	2090	2100
TGCAGTCCGG	AGCTACTCAT	CAGGTGGTAT	ACTGGATCAT	TCCTTTGCC	ATGGCTTAGA
2110	2120	2130	2140	2150	2160
AACCACTACG	TCAAGAAGGA	CAGGAAATGG	TTCCAGGAAC	CATACCGTA	CCCCAAGCAT
2170	2180	2190	2200	2210	2220
CTTGAAACCC	ATCCAGAGCT	CGCAGATCAA	GCATGGCTTT	ACAAATCTGT	TCTAGAAATT
2230	2240	2250	2260	2270	2280
TGCAGATACT	GGGTAGAGCT	AAGATATTCC	CTCATCCAGC	TCCTTTACGA	CTGCATGTC
2290	2300	2310	2320	2330	2340
CAAAACGTGG	TCGATGGTAT	GCCACTTGCC	AGATCTATGC	TCTTGACCGA	TACTGAGGAT
2350	2360	2370	2380	2390	2400
ACGACCTTCT	TCAATGAGAG	CCAAAAGTTC	CTCGATAACC	AATATATGGC	TGGTGACGAC
2410	2420	2430	2440	2450	2460
ATCCTTGTAG	CACCCATCCT	CCACAGCCGT	AACGAGGTTC	CGGGAGAGAA	CAGAGATGTC
2470	2480	2490	2500	2510	2520
TATCTCCCTC	TATTCCACAC	CTGGTACCCC	TCAAACCTGA	GACCGTGGGA	CGATCAGGGAA
2530	2540	2550	2560	2570	2580
GTCGCTTTAG	GGAATCCTGT	CGAAGGTGGC	AGCGTTATCA	ACTACACTGC	CAGGATTGTT
2590	2600	2610	2620	2630	2640
GCCCCAGAGG	ATTATAATCT	CTTCCACAAAC	GTGGTGCCGG	TCTACATCAG	AGAGGGTGCC
2650	2660	2670	2680	2690	2700
ATCATTCCGC	AAATTCAAGT	ACGCCAGTGG	ATTGGCGAAG	GAGGGCCTAA	TCCCATCAAG
2710	2720	2730	2740	2750	2760
TTCAAATATCT	ACCCCTGGAAA	GGACAAGGGAG	TATGTGACGT	ACCTTGATGA	TGGTGTTAGC
2770	2780	2790	2800	2810	2820
CGCGATAGTG	CACCAAGATGA	CCTCCCGCAG	TACCGCGAGG	CCTATGAGCA	AGCGAAGGTC
2830	2840	2850	2860	2870	2880

FIGURE 10 CONTINUED

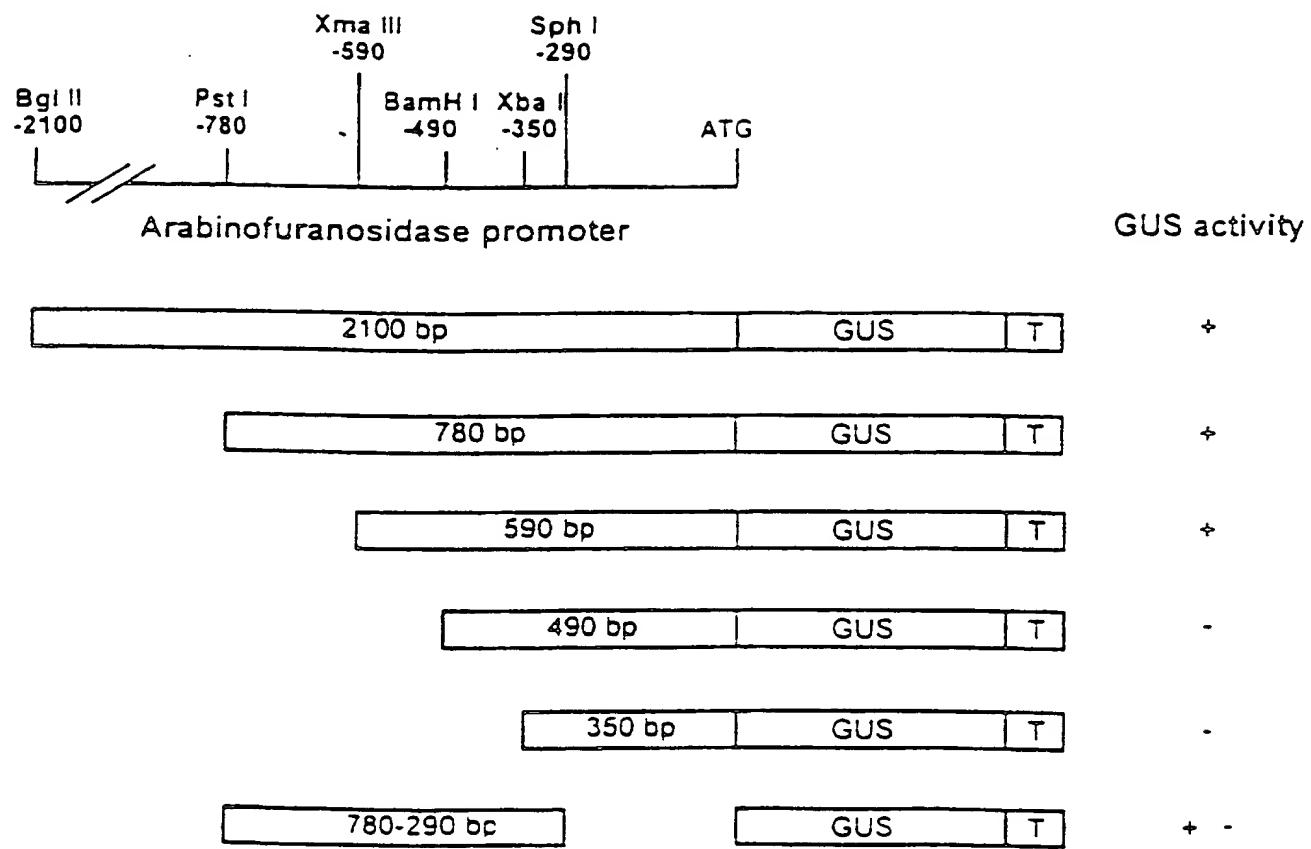
GAAGGCCAAAG ACGTCCAGAA GCAACTTGCG GTCATTCAAG GGAATAAGAC TAATGACTTC
2890 2900 2910 2920 2930 2940
TCCGCCTCG GGATTGATAA GGAGGCCAAAG GGTTATCACC GCAAAGTTTC TATCAAACAG
2950 2960 2970 2980 2990 3000
GAGTCAAAAG ACAAGACCCG TACTGTCAACC ATTGAGCCAA AACACAACGG ATACGACCCC
3010 3020 3030 3040 3050 3060
TCTAAGGAAG TTGGTAATTAA TTATACCATC ATTCTTTGGT ACGCACCGGG CTTTGACGGC
3070 3080 3090 3100 3110 3120
AGCATCGTGG ATGTGAGCCA GGCGACCGTG AACATCGAGG GCGGGGTGGA ATGCGAAATT
3130 3140 3150 3160 3170 3180
TTCAAGAACCA CGGGCTTGCA TACGGTTGTA GTCAACGTGA AAGAGGTGAT CGGTACCCACA
3190 3200 3210
AAGTCCGTCA AGATCACTTG CACTACCGCT TAG

FIG. 11



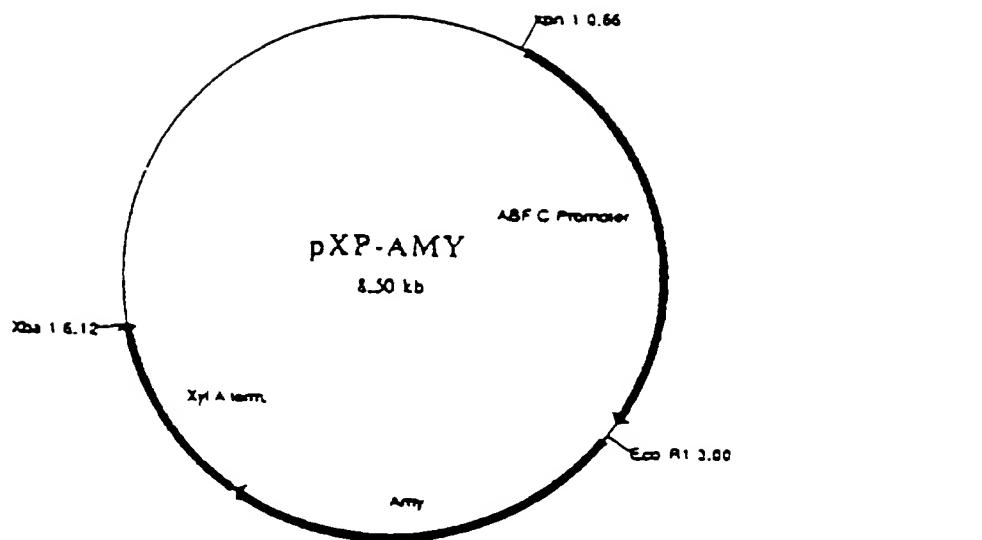
28/38

FIG. 12



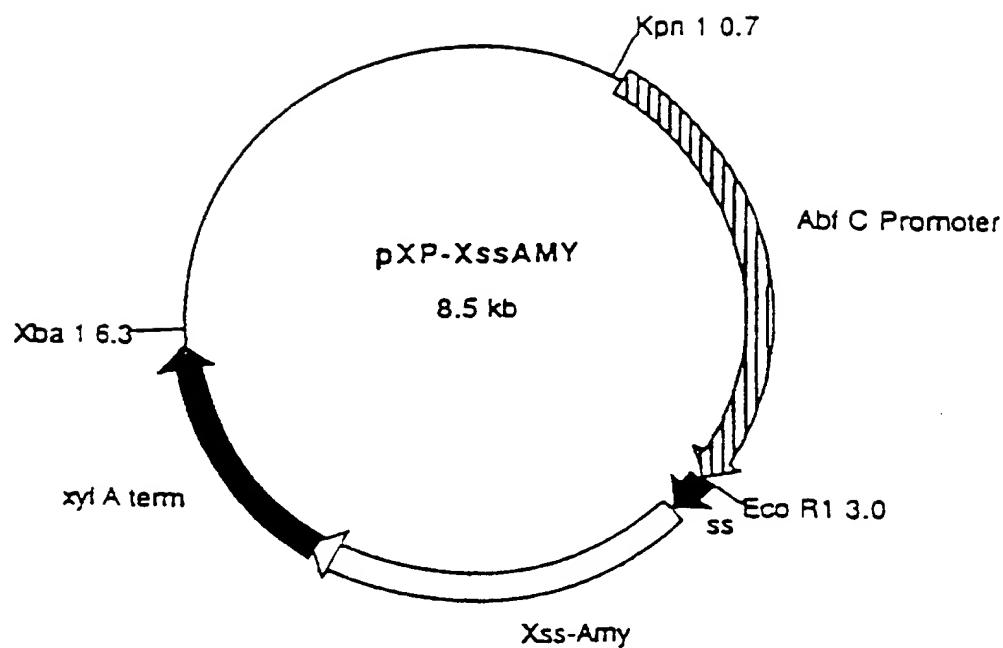
29/38

FIG. 13



30/38

FIG. 14



31/38

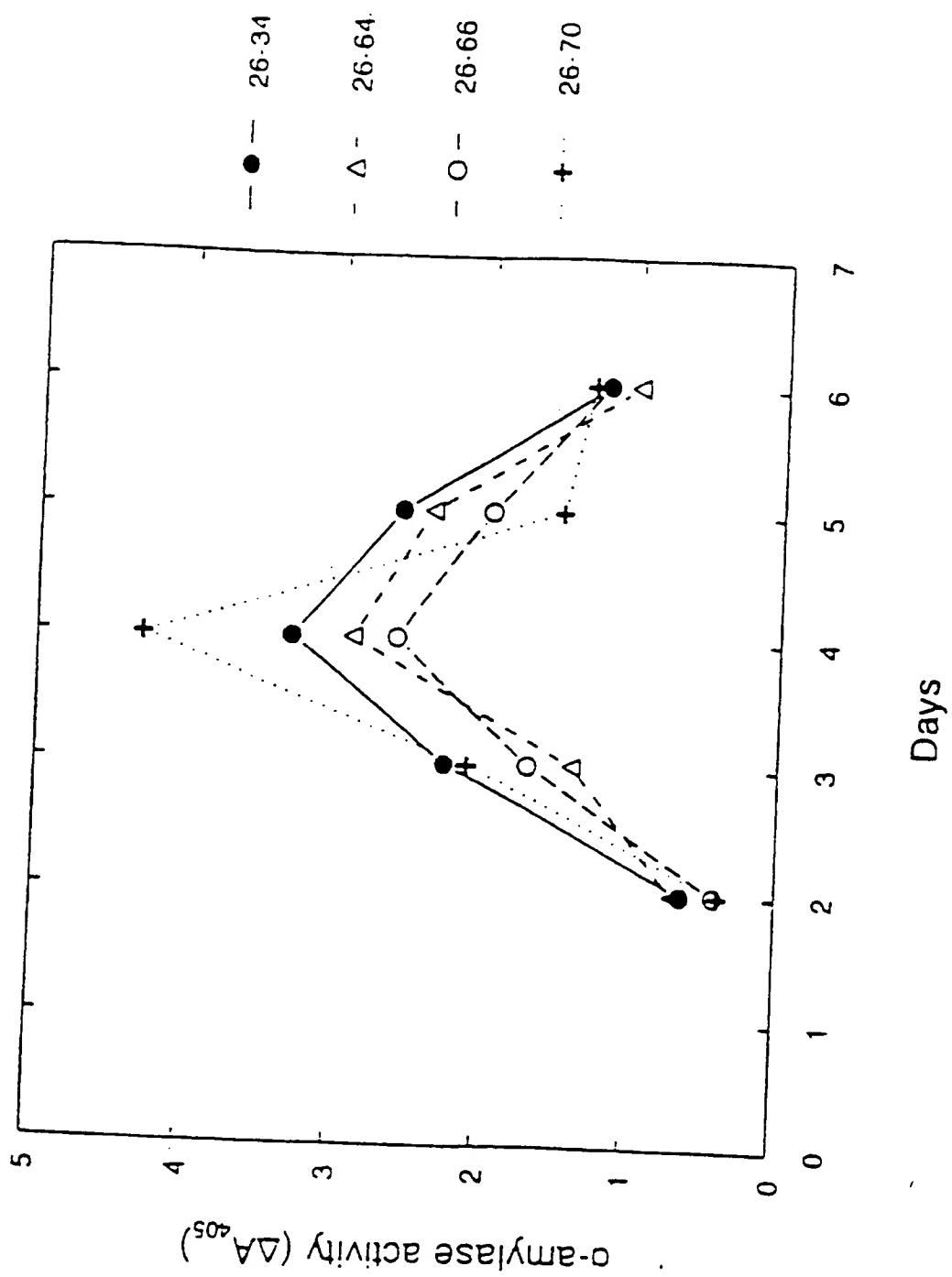


FIG. 15

32/38

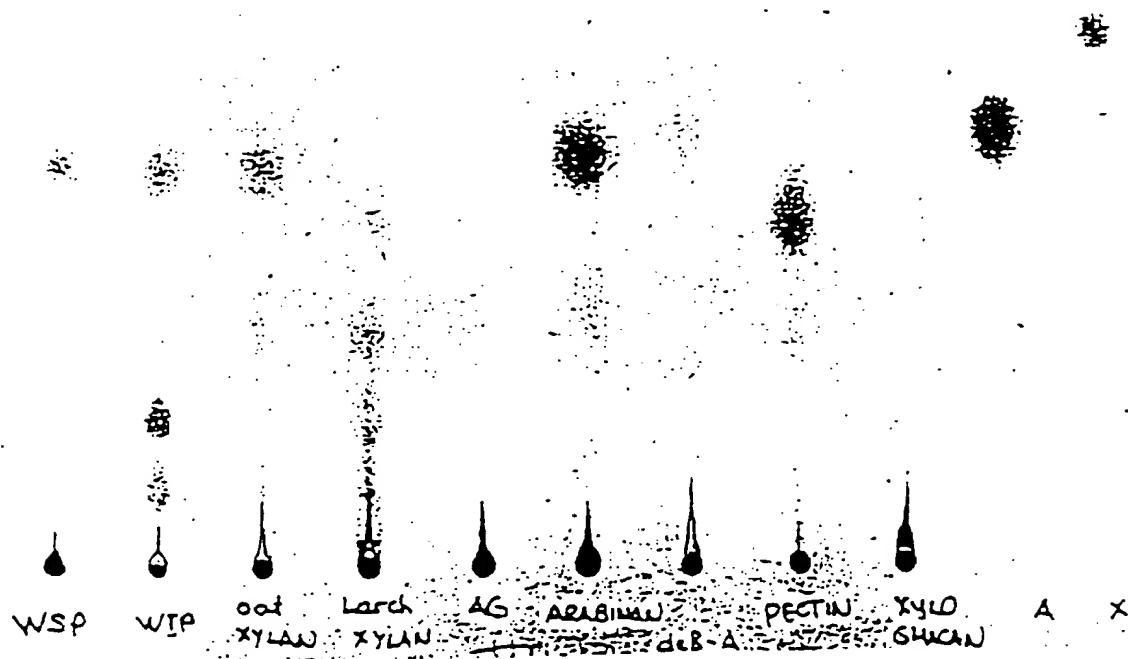


FIG. 16

33/38

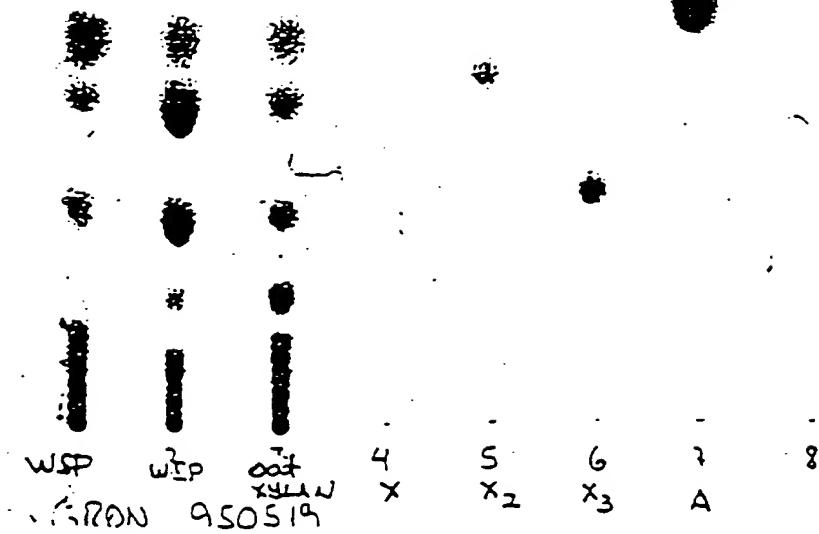


FIG. 17

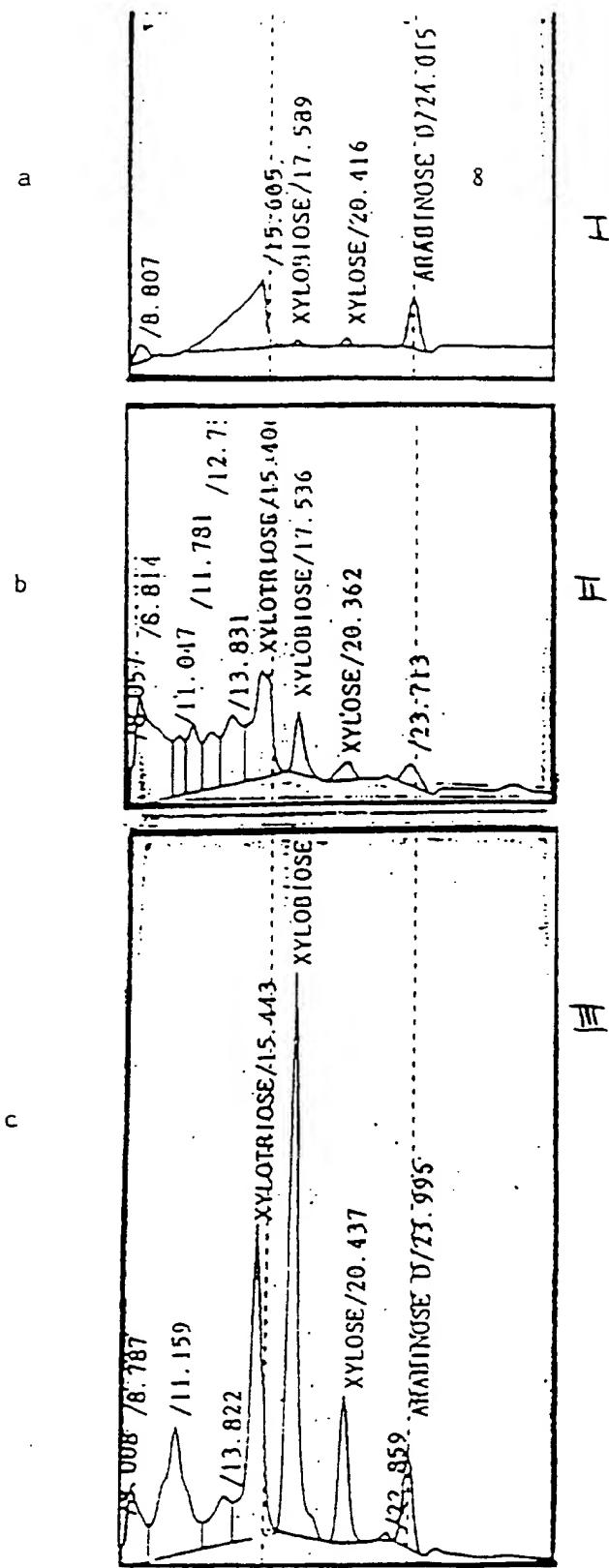


FIG. 18

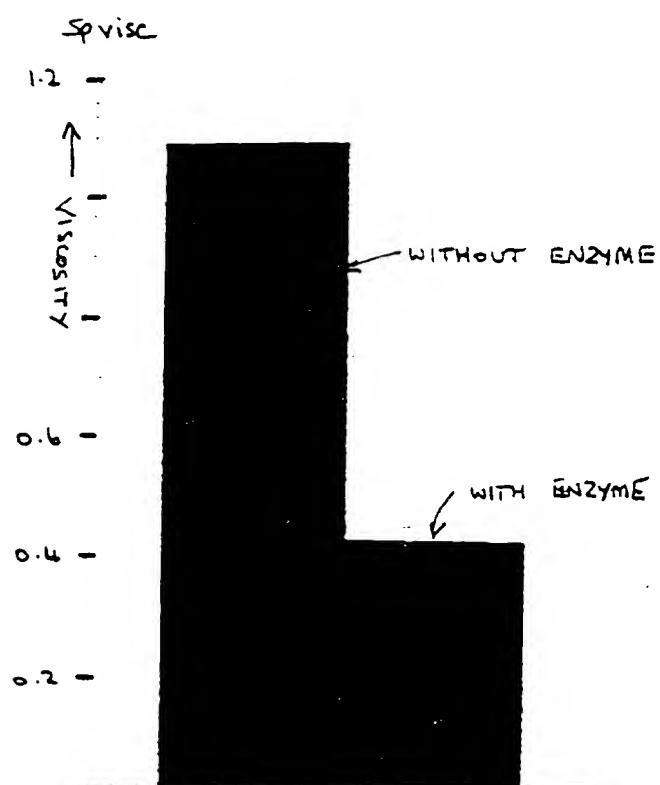


FIG. 19

Fig 20

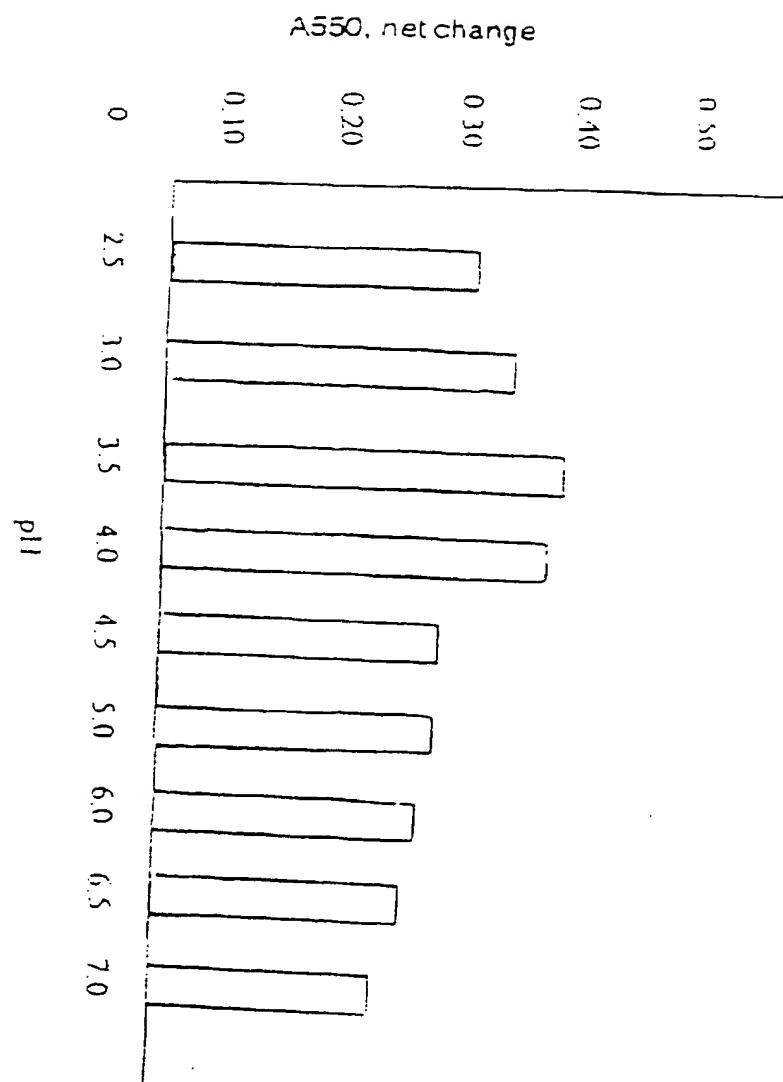
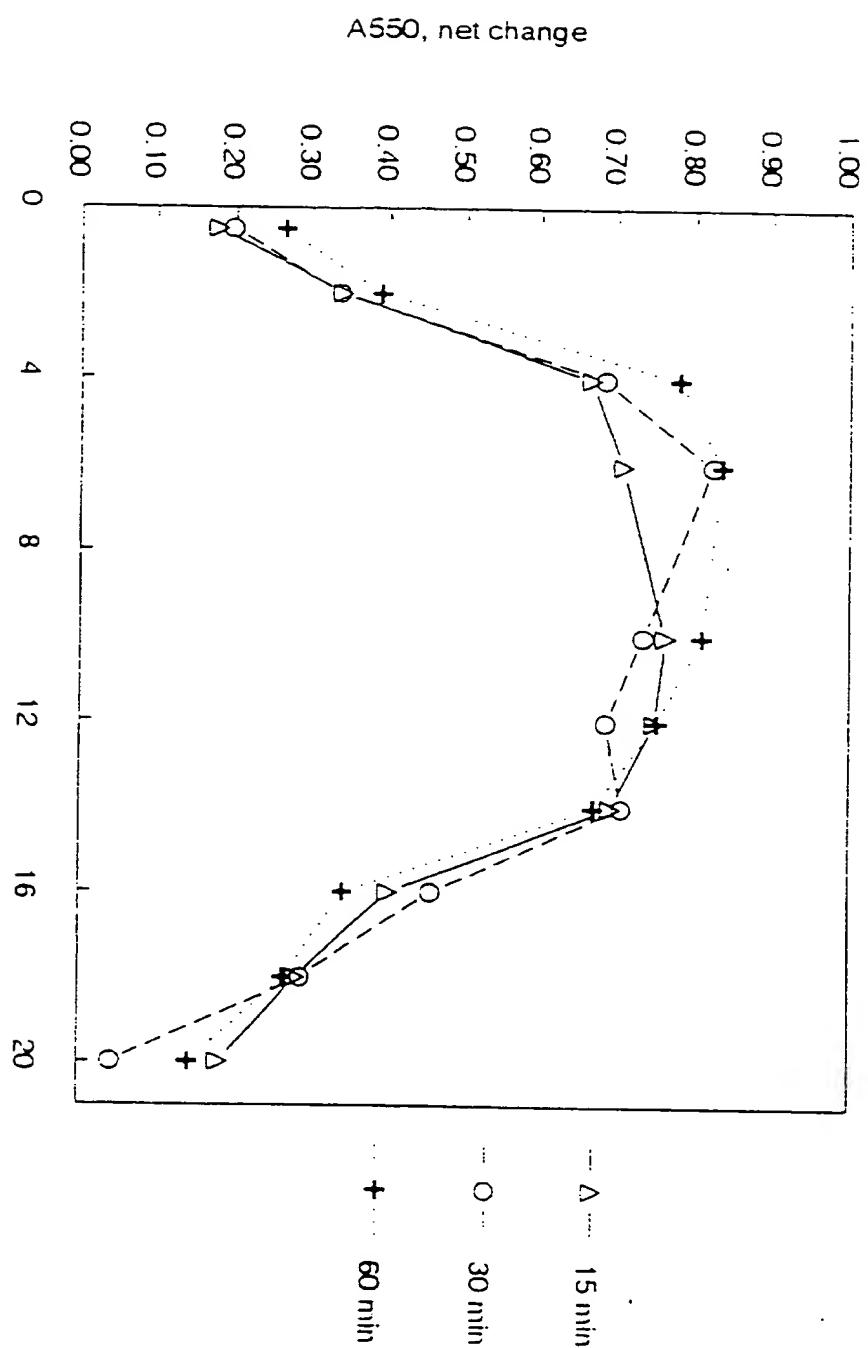


Fig 21



A550, net change

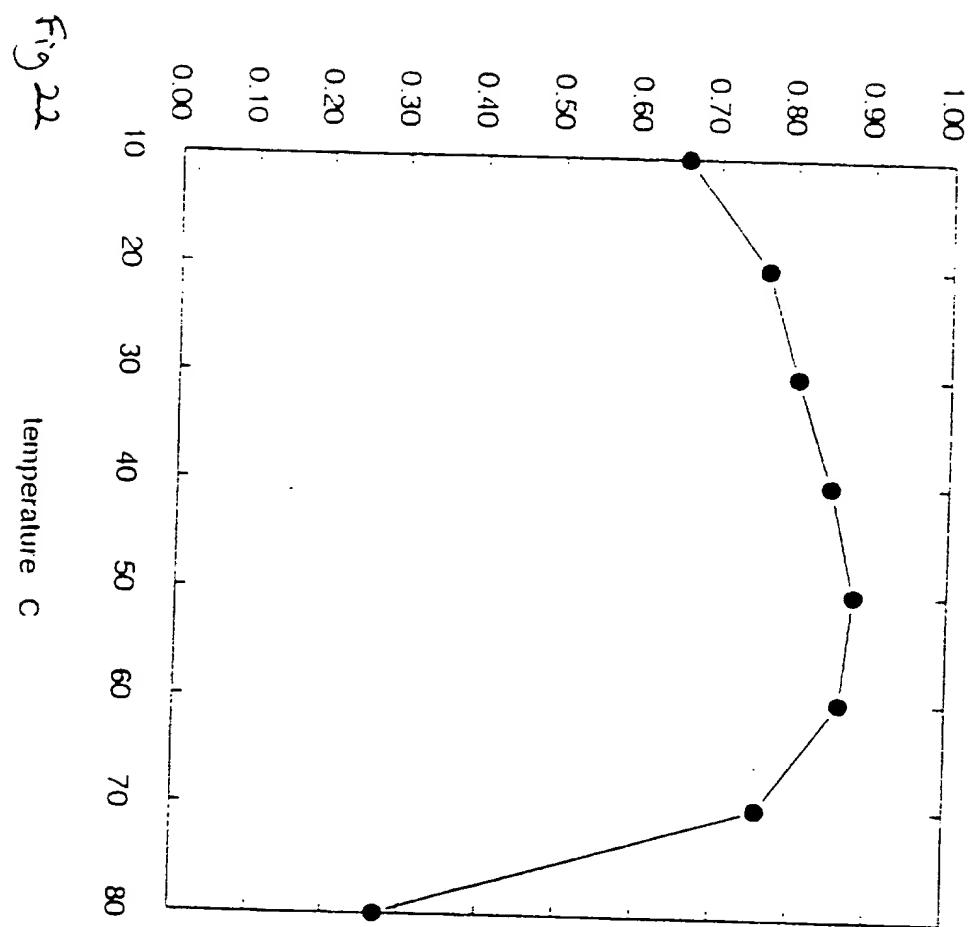


Fig 22

temperature C

A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 C12N15/56	C12N15/80	C12N1/15	C12P19/04	A01H5/00
A61K38/47	A23L1/29			

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C12N C12P A01H A61K A23L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	APPLIED MICROBIOLOGY AND BIOTECHNOLOGY, vol. 35, no. 6, September 1991, pages 753-758, XP000574930 KORMELINK, F.J.M., ET AL.: "Purification and characterization of a (1,4)-beta-D-arabinoxyran arabinofuranohydrolase from Aspergillus awamori" cited in the application see the whole document ---	1,2,26, 29,31
P,X	WO,A,96 06935 (GIST BROCADES BV ;WOUW MONIQUE JOSINA ANDREA V D (NL); OOIJEN ALBE) 7 March 1996 see the whole document ---	1-19, 21-25, 29-31 -/-

Further documents are listed in the continuation of box C.

Patent family members are listed in annex.

* Special categories of cited documents :

- 'A' document defining the general state of the art which is not considered to be of particular relevance
- 'E' earlier document but published on or after the international filing date
- 'L' document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- 'O' document referring to an oral disclosure, use, exhibition or other means
- 'P' document published prior to the international filing date but later than the priority date claimed

- 'T' later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- 'X' document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- 'Y' document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
- '&' document member of the same patent family

1

Date of the actual completion of the international search	Date of mailing of the international search report
19 July 1996	29.07.96

Name and mailing address of the ISA
European Patent Office, P.B. 5818 Patentstaan 2
NL - 2280 HV Rijswijk
Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl.
Fax (+ 31-70) 340-3016

Authorized officer

Maddox, A

INTERNATIONAL SEARCH REPORT

International Application No
PCT/EP 96/01009

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	EP,A,0 463 706 (GIST BROCADES NV) 2 January 1992 see the whole document ---	1-31
A	WO,A,90 01059 (UNIV CALIFORNIA) 8 February 1990 see the whole document ---	1-31
A	BIOTECHNOLOGY, vol. 13, no. 1, January 1995, page 63 XP002008876 HERBERS, K., ET AL.: "A thermostable xylanase from Clostridium thermocellum expressed at high levels in the apoplast of transgenic tobacco has no detrimental effects and is easily purified" see the whole document ---	20
A	EP,A,0 479 359 (GIST BROCADES NV ;MOGEN INT (NL)) 8 April 1992 see page 24, line 25 - line 32 ---	20
A	EP,A,0 507 369 (GIST BROCADES NV) 7 October 1992 see claims 11-13 ---	25-27,30
A	WO,A,92 10945 (ENZYMATIX LTD ;BIOTAL LTD (GB)) 9 July 1992 see the whole document ---	25-27,30
A	EP,A,0 506 190 (GIST BROCADES NV) 30 September 1992 cited in the application see claim 36 ---	25-27,30
A	BIOCHEMICAL JOURNAL, vol. 272, no. 2, 1990, pages 369-376, XP000576098 KELLETT, L.E., ET AL.: "Xylanase B and arabinofuranosidase from Pseudomonas fluorescens subsp. cellulosa contain identical cellulose-binding domains and are encoded by adjacent genes" see the whole document -----	1-6

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/EP 96/01009

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO-A-9606935	07-03-96	AU-B- 3387795	22-03-96
EP-A-0463706	02-01-92	AT-T- 124844 AU-B- 654147 AU-B- 8318691 AU-B- 647170 AU-B- 8320591 CA-A- 2066734 CA-A- 2067329 DE-D- 69111148 DE-T- 69111148 EP-A- 0468596 JP-T- 5501657 JP-T- 5500907 WO-A- 9201389 WO-A- 9201793 US-A- 5358864	15-07-95 27-10-94 18-02-92 17-03-94 18-02-92 25-01-92 25-01-92 17-08-95 04-01-96 29-01-92 02-04-93 25-02-93 06-02-92 06-02-92 25-10-94
WO-A-9001059	08-02-90	US-A- 4954447	04-09-90
EP-A-0479359	08-04-92	AU-B- 656920 AU-B- 8651491 CA-A- 2072656 JP-T- 5502591 WO-A- 9205259 NZ-A- 239789	23-02-95 15-04-92 14-03-92 13-05-93 02-04-92 27-06-94
EP-A-0507369	07-10-92	CA-A- 2063352 JP-A- 5123173 US-A- 5426043	19-09-92 21-05-93 20-06-95
WO-A-9210945	09-07-92	AT-T- 126972 AU-B- 9062291 CA-A- 2098708 DE-D- 69112643 DE-T- 69112643 EP-A- 0563133 ES-T- 2079173 US-A- 5432074	15-09-95 22-07-92 18-06-92 05-10-95 07-03-96 06-10-93 01-01-96 11-07-95

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/EP 96/01009

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP-A-0506190	30-09-92	AU-B- 651194	14-07-94
		AU-B- 1575592	02-11-92
		CA-A- 2084056	28-09-92
		JP-T- 6500022	06-01-94
		WO-A- 9217592	15-10-92

THIS PAGE BLANK (USPTO)